An Empirical Analysis of Survey-Based Macroeconomic Forecasts and Uncertainty

Inauguraldissertation

zur Erlangung des akademischen Grades
Doctor Rerum Politicarum

an der
Fakultät für Wirtschafts- und Sozialwissenschaften
der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Alexander Glas

Heidelberg
eingereicht im Juli 2018

Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisor, Christian Conrad, for his stimulating advice throughout my academic years. I greatly benefited from his valuable and encouraging feedback on my research. I am also very thankful to Jonas Dovern for his immediate readiness to be a member of the thesis committee.

The quality of this thesis has been improved considerably by the fruitful collaboration with my co-authors Christian Conrad and Matthias Hartmann. I am deeply indebted to their intensive work on three out of four research papers upon which this thesis is based. I would also like to thank Julian Donaubauer, Michael Hübler, Birgit Meyer and Peter Nunnenkamp for the inspiring cooperation on several other projects.

I immsensely enjoyed the working atmosphere at the Professorship of Empirical Economics at Heidelberg University and the collaboration with my colleagues Onno Kleen, Fabian Krüger, Karin Stürmer and Michael Stollenwerk. My research has also been improved by participating in the internal and Macro & Econometrics seminars at the Alfred-Weber-Institute as well as the workshops organized by the HeiKaMEtrics network. In particular, I would like to thank Carsten Bormann, Rebekka Buse, Thomas Eife, Zeno Enders, Alexander Jordan, Lennart Kaplan, Sarah Langlotz, Sebastian Lerch, Chong Liang, Katharina Richert, Jörg Rieger, Melanie Schienle, Sandra Schmidt, Carsten Trenkler, David Vespermann and Christopher Zuber. In addition, I benefited from the comments and suggestions by the participants of numerous international academic conferences.

Moreover, I am grateful to the lecturers and fellow teaching assistants that I have had the pleasure to work with at Heidelberg University. Among those not mentioned above are Melanie Arntz, Anna Busse, Sebastian Butschek, Philipp Greuell, Christian Hengesbach, Franziska Lembcke, Andreas Lindenblatt, Yuhong Mangels, Angelika Müller, André Nolte, Nestor Parolya, Lora Pavlova, Hannes Vetter and Arne Warnke. I also thank the participants of my tutorials for turning teaching into such a fun and rewarding experience.

Finally, I am extremely grateful for the steady support of my parents over the last years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Survey-Based Macroeconomic Forecasts

Surveys of macroeconomic expectations have become increasingly popular in the forecasting literature and provide a convenient alternative to forecasts from time series models. Three distinct types of survey participants can be distinguished. The first group includes expert forecasters from financial and research institutes. Prominent examples are the Survey of Professional Forecasters (SPF) operated by the European Central Bank (ECB) or the Federal Reserve Bank of Philadelphia (FED) and The Bank of England's Survey of External Forecasters. The second group focuses on consumer predictions. Two of the most popular examples from this group are the University of Michigan's Survey of Consumers and the Federal Reserve Bank of New York's Survey of Consumer Expectations. A disadvantage of using these surveys is that consumers have been shown to frequently misinterpret economic and/or mathematical concepts such as 'inflation' or 'percentage change' (e.g., Bruine de Bruin et al., 2011, 2012). The third group consists of business and firm surveys such as the Business Outlook Survey operated by the Federal Reserve Bank of Philadelphia. In this thesis, we focus on the predictions from expert forecasters since professionals should have a clear understanding of economic concepts and relationships and thus provide the most reliable predictions.

Predictions by professional forecasters can be used to assess the overall economic outlook as well as the credibility of monetary policy. For example, the ECB frequently highlights the necessity of stable inflation expectations as a prerequisite for the conduct of successful monetary policy. In particular, the ECB emphasized the importance of well-anchored inflation expectations for stabilizing inflation rates in the Euro area in the aftermath of the Great Recession and the Sovereign Debt Crisis:

> "*The accommodative monetary policy stance will underpin the firm anchoring of medium to long-term inflation expectations, in line with the Governing Council's aim of achieving inflation rates below, but close to, 2%.*"
>
> (ECB, 2014a, p. 5)

Point forecasts from expert surveys provide a readily available option to gauge whether long-term inflation expectations are anchored. Several desirable properties of survey-based point predictions have been documented in the literature. For example, Dräger et al. (2016) find that the FED-SPF expectations are consistent with theoretical economic relationships such as the Fisher equation, the Taylor rule or the Phillips curve. With respect to forecast accuracy, Ang et al. (2007) show that inflation forecasts from the FED-SPF and the Livingston survey outperform predictions from various econometric models. Faust and Wright (2009) document similar evidence for the Greenbook forecasts. In Chapter 2, we contribute to the literature by investigating whether survey forecasts improve predictions of future stock return volatility.

The survey data used throughout this thesis are taken from the SPF of both the ECB and the FED. The SPF is a quarterly survey of macroeconomic expectations and includes participants from the financial services industry as well as from research institutes. The ECB-SPF has been conducted since 1999Q1 and asks participants to report their expectations for the inflation rate, real GDP growth and the unemplyoment rate in the Euro area. The FED-SPF has been conducted in consecutive quarters since 1968Q4 and is thus available for a considerably longer time period. Before 1990, the FED-SPF was operated by the American Statistical Association and the National Bureau of Economic Research. In 1990, the FED took over the responsibilities for the SPF. In addition to the outcome variables covered by the ECB-SPF, the questionnaire of the FED-SPF asks respondents to report predictions for several other measures of economic activity, such as industrial production or housing starts, as well as various financial variables. Both incarnations of the SPF ask panelists for both point predictions and subjective probability distributions in the form of histogram forecasts for a variety of forecast horizons. The empirical analysis presented throughout this thesis employs both types of forecasts for distinct outcome variables and forecast horizons. Further details will be explained in subsequent chapters.

1.2 Remarks on Macroeconomic Uncertainty

In his Nobel Price acceptance speech, Friedman (1977) emphasized the negative impact of heightened inflation uncertainty on economic growth and the level of employment. In an increasingly uncertain environment, economic agents may take a 'wait-and-see' approach and postpone investment and/or saving decisions until the uncertainty is resolved. Empirical research confirms the negative impact of increased macroeconomic uncertainty on economicy activity (e.g., Bloom et al., 2014; Jurado et al., 2015; Meinen and Röhe, 2017). Despite this, few institutions reported the uncertainty associated with their predictions before the outbreak of the Great Recession. One of the few exceptions was the Bank of England, which enhanced its macroeconomic projections with uncertainty bands since 1997. The failure of many professional forecasters to predict the outbreak of the Great Recession has put the sole reliance on point predictions into question. This led to an increasing tendency to report quantifications of the uncertainty associated with the point forecasts. Chapters 3 through 5 are primarily concerned with the measurement of macroeconomic uncertainty as well as its covariates.

One of the primary challenges of studying uncertainty is its elusive nature. Uncertainty is a latent variable and therefore inherently unobserved. One of the most popular ways to quantify uncertainty ex-ante is to use either point or histogram forecasts from professionals (Giordani und Söderlind, 2003). Due to unavailability of histogram forecasts, early research in this field focused on the dispersion of the first moments, i.e., the individual point predictions. The cross-sectional variance of point forecasts is commonly referred to as 'forecaster disagreement'. More recently, researchers have begun to use the second moments from the subjective probability distributions to compute indicators of uncertainty. In particular, the average variance of the individual histogram forecasts is another popular indicator of uncertainty ('average uncertainty'). A large body of literature investigates the question of whether disagreement is a good proxy for average uncertainty (e.g., Lahiri and Sheng, 2010). Theoretically, there is no conclusive answer to this question so far (Zarnowitz and Lambros, 1987). Empirically, Rich and Tracy (2010) and Abel et al. (2016), among others, show that disagreement is only weakly correlated with average uncertainty in the SPF data.

In Chapters 3 and 4 of this thesis, we provide a comprehensive analysis of the relationship between disagreement and uncertainty as well as their respective covariates. In Chapter 3, special attention is drawn to the role of discretionary monetary policy for the emergence of heightened inflation uncertainty since the outbreak of the Great Recession. During this period, the monetary policy of the ECB has been characterized by extremely low interest rates close to the zero lower bound as well as large-scale asset purchases. Such deviations from a rules-based monetary policy may have contributed to the rise in inflation uncertainty. In Chapter 4, we provide a broader analysis of the relationship between uncertainty and disagreement that accounts for distinct outcome variables as well as technical assumptions regarding the construction of such measures.

Several studies document desirable properties of uncertainty proxies derived from the histogram forecasts. For example, Lahiri and Sheng (2010) show that average uncertainty in the FED-SPF increases with the forecast horizon. However, the calibration of the histogram forecasts has been questioned recently by Clements (2016), who shows that short-horizon survey predictions do not outperform unconditional density forecasts based on the past level of uncertainty.[1] His analysis is part of a wider branch of literature that focuses on the observation that the ex-ante variance frequently deviates from the average squared ex-post forecast error. Empirically, a misalignment of ex-ante and ex-post uncertainty in the SPF data has been documented by Giordani and Söderlind (2003, 2006), Kenny et al. (2014, 2015) and Clements (2014). This is frequently referred to as evidence of forecaster 'over- or underconfidence'.

In Chapter 5, we contribute to the literature by analyzing the relationship between the variance misalignment and the coarseness of survey predictions. In particular, we show that the participants of the SPF consist of two distinct groups, which we refer to as 'rounders' and 'non-rounders', who significantly differ in terms of the degree of rounding of their predictions as well as the degree of the variance misalignment. We propose that a more accurate assessment of uncertainty is possible by focusing on the group of non-rounders when calculating uncertainty. Moreover, we discuss to what extent the rounding choices of survey participants may be related to different ways of expectation formation, e.g., the role of judgment versus the use of (econometric) models.

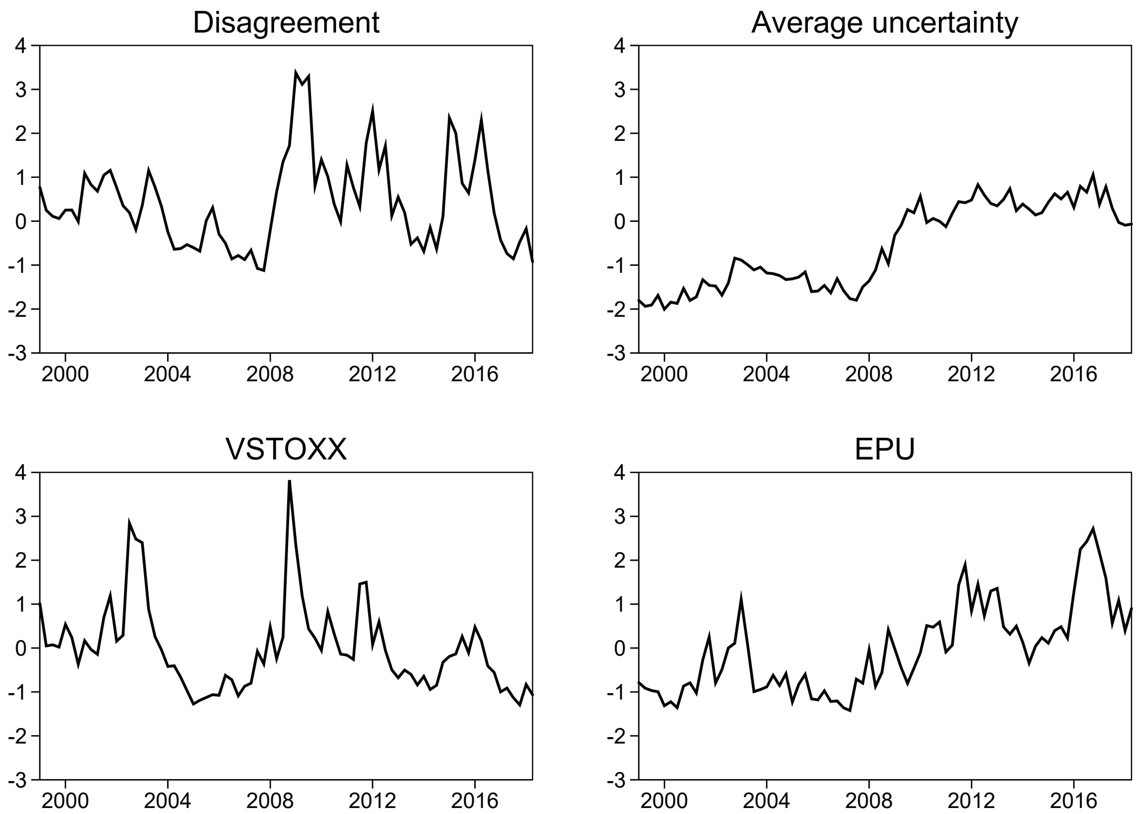
It is important to highlight that various other quantifications of uncertainty besides survey-based measures have been proposed in the literature. These measures can be broadly categorized as ex-post and ex-ante indicators. Ex-post measures are usually based on realizations and comprise, e.g., the realized standard deviation of the outcome variable (Klein, 1977) or squared prediction errors (Barron et al., 1998). Another popular indicator of uncertainty is the Economic Policy Uncertainty index (EPU) developed by Baker et al. (2016), which measures the number of newspaper articles related to economic uncertainty. Ex-ante measures include forecasts of the conditional variance of the outcome variable from time series models. For example, the conditional variance may be modeled as a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) process (Engle, 1983). A forward-looking measure of the uncertainty on U.S. financial markets is the Volatility Index (VIX) developed by the Chicago Board Options Exchange (CBOE). This index captures the expected stock market volatility implied by S&P 500 index options. The European counterpart of the VIX is the EURO STOXX 50 Volatility

[1] The terms 'histogram forecast' and 'density forecast' are used synonymously throughout.

index (VSTOXX).

To highlight the differences of various popular uncertainty proxies, Figure 1.1 depicts the quarterly time series of forecaster disagreement and average uncertainty based on one-year-ahead inflation forecasts from the ECB-SPF as well as the quarterly averages across the VSTOXX and EPU indices for the Euro area during the period 1999Q1–2018Q2. Each series is standardized to have zero mean and variance one. The measures depicted in Figure 1.1 will be discussed in more detail throughout the remaining chapters. Thus, we refrain from providing further details here.

Figure 1.1: Popular indicators of macroeconomic uncertainty



Notes: The plots depict the standardized time series of four popular indicators of macroeconomic uncertainty: Forecaster disagreement and average uncertainty based on one-year-ahead inflation forecasts from the ECB-SPF, the EURO STOXX 50 Volatility index as well as the Economic Uncertainty index. The sample period is 1999Q1–2018Q2.

The evidence from Figure 1.1 shows that the individual measures considerably deviate during specific episodes, in particular since the outbreak of the Great Recession. While forecaster disagreement exhibits subtantial fluctuations, average inflation uncertainty has experienced a sustained increase since around 2008. A similar divergence of the EPU and VIX indices is visible in recent years. In addition, disagreement appears to track the evolution of the VSTOXX index, whereas average uncertainty appears to co-move more closely with the EPU index. These differences highlight the need to account for the conceptual differences of distinct uncertainty proxies. To address this issue, Chapters 2 through 4 include comparisons of survey-based measures of uncertainty to other popular indicators of macroeconomic uncertainty.

1.3 Outline of the Thesis

This section provides an outline of the remainder of this thesis, which consists of four self-contained chapters that can be read independently of each other. The ordering of chapters mirrors the structure of the discussion from the previous subsections. Chapter 2 is joint work with my supervisor, Prof. Dr. Christian Conrad. Chapters 3 and 5 are joint projects with one my former colleagues, Dr. Matthias Hartmann. A slightly different version of the analysis from Chapter 3 has been published as "Inflation uncertainty, disagreement and monetary policy: Evidence from the ECB Survey of Professional Forecasters" in a special issue of the *Journal of Empirical Finance* on "The Euro Zone in Crisis" (cf. Glas and Hartmann, 2016). Chapter 4 is based on a single-authored paper. Below, we motivate the analysis in the subsequent chapters and briefly summarize the main results.

'Déjà Vol' Revisited

In Chapter 2, we investigate the question of whether macroeconomic variables contain information about future stock volatility beyond that contained in past volatility. Among other models of asset valuation, the present value models of Campbell (1991) and Campbell and Shiller (1988) provide a theoretical justification for counter-cyclical behavior of stock volatility. Nonetheless, the empirical evidence for a link between volatility and macroeconomic conditions is rather weak. For example, Schwert (1989) and Paye (2012) show that macroeconomic variables rarely improve predictions of future stock market volatility when conditioning on past volatility.

Using data from the FED-SPF, we contribute to the literature by providing strong evidence that macroeconomic variables do successfully predict quarterly stock volatility. In contrast to the previous literature which focuses almost exclusively on the realizations of macroeconomic variables, we show that corresponding survey forecasts are usually more informative. Moreover, the release schedule of quarterly macroeconomic figures rules out the estimation of predictive regressions with realizations in real-time. Since they do not account for this publication lag, the results from previous studies are thus of limited use in practice. In addition, previous research mostly considers the stock market as a whole. However, stock volatility in some industries may be more closely linked to macroeconomic conditions than in others. Therefore, we analyze both the aggregate stock market as well as 5 and 49 industry portfolios. Finally, we argue that predictive regressions based on realized volatility suffer from measurement error, which may mask a potentially existing effect of macroeconomic conditions on future stock volatility. To mitigate this problem, we follow Conrad and Schienle (2018) and replace realized volatility with 'volatility-adjusted realized volatility', which is a more precise proxy for the latent long-term volatility component that is driven by macroeconomic conditions.

Our results can be summarized as follows: First, we find that forecasts of real GDP and industrial production growth predict volatility in a cross-section of 49 industry portfolios. We show that the expectation of higher growth rates is associated with lower stock volatility. Our results are in line with both counter-cyclical volatility in dividend news as well as in expected returns. In contrast, realizations of GDP and industrial production growth are essentially insignificant predictors in all industries. Second, inflation forecasts predict higher or lower stock volatility depending on the state of the economy and the stance of monetary policy. Third, forecasts of higher unemployment rates are good news for stocks during expansions and go along with lower stock volatility. The results from predictive regressions with 'volatility-adjusted realized volatility' strengthen the evidence in favor of the predictive ability of macroeconomic variables. In particular, the realizations gain significance in this case. However, they are still outperformed by the forecasts. In summary, we provide much more optimistic results regarding the predictive power of macroeconomic variables for stock volatility than the previous literature. The results from our analysis hold in- as well as out-of-sample and pass various robustness checks.

Inflation Uncertainty, Disagreement and Monetary Policy

After demonstrating the predictive power of survey forecasts for future stock volatility in Chapter 2 we next turn to the analysis of macroeconomic uncertainty. In Chapter 3, we focus on the uncertainty associated with expected inflation rates in the Euro area. Though survey-based measures of inflation uncertainty such as the average variance from a cross-section of density forecasts are often regarded as one of the most reliable ways to quantify uncertainty (Bachmann et al., 2013; Clements, 2014), many surveys do not elicit density forecasts. In such cases, the cross-sectional variance of point forecasts (i.e., disagreement) is often used as a proxy variable for inflation uncertainty. However, Zarnowitz and Lambros (1987) describe several cases in which the two measures may deviate. For the FED-SPF, Giordani and Söderlind (2003) as well as Rich and Tracy (2010) document that disagreement is only weakly correlated with measures of inflation uncertainty derived from density forecasts.

In Chapter 3, we analyze the covariates of average inflation uncertainty and forecaster disagreement based on data from the ECB-SPF. In particular, we are the first to provide a systematic analysis of the role of discretionary monetary policy for the emergence of inflation uncertainty in the Euro area. Employing the forecast error component model of Lahiri and Sheng (2010), we conduct a variance decomposition which isolates disagreement as one of two components of average inflation uncertainty. The second component is the perceived variance of the shocks that occur during the period after forecasters report their predictions until the realization of the inflation rate. We empirically confirm that disagreement is an incomplete approximation to overall uncertainty. Both measures are associated with macroeconomic conditions and indicators of monetary policy, but the relations differ qualitatively. In particular, we show that average inflation uncertainty is higher during periods of expansionary monetary policy, whereas disagreement rises during periods when monetary policy is discretionary. Moreover, the difference between average inflation uncertainty and disagreement increases with the forecast horizon and during periods when monetary policy is unduly expansive. Our findings suggest that an important reason for the steady increase in average inflation uncertainty since the beginning of the Great Recession is the sustained period of expansionary monetary policy by the ECB. This influence is not detected if disagreement alone is used as an indicator of inflation uncertainty. Therefore, conclusions based on disagreement as a single indicator of ex-ante uncertainty are incomplete and potentially misleading.

Five Dimensions of the Uncertainty-Disagreement Linkage

The analysis from Chapter 3 is based on a very specific setting. In particular, we employ variance-based indicators of uncertainty and disagreement and use the histogram means to compute the latter. Moreover, we fit generalized beta distributions to the individual histograms and focus on predictions for the inflation rate. Chapter 4 provides a broader analysis of the relationship between forecaster disagreement and macroeconomic uncertainty in the Euro area. Using data from the ECB-SPF for the period 1999Q1–2018Q2, we account for several aspects of the uncertainty-disagreement linkage that are frequently neglected or insufficiently addressed in the related literature. First, we consider distinct statistics that differ in terms of their robustness to extreme observations. Most studies rely on variance-based indicators, which may be strongly affected by outliers (Lahiri and Sheng, 2010). This is practically relevant because the cross-section of survey participants is relatively small in most cases. Thus, a few extreme predictions may falsely indicate a large degree of disagreement among survey participants. Second, both the point forecasts and the histogram means are frequently used to calculate forecaster disagreement. However, the two series tend to deviate in practice (Clements, 2010, 2012). The choice between the two may thus have a considerable impact on the observed level of disagreement. In order to assess how this choice affects the link between uncertainty and disagreement, we consider both. Third, most studies rely on a single assumption regarding the distribution of the probability mass within the distinct outcome intervals of the survey questionnaire. In contrast, we account for various popular assumptions. As discussed by Giordani and Söderlind (2003), this choice affects the observed level of uncertainty. Fourth, most studies that analyze the SPF data focus on either the inflation rate (D'Amico and Orphanides, 2008; Glas and Hartmann, 2016), output growth (Clements, 2011), or both (Zarnowitz and Lambros, 1987; Giordani and Söderlind, 2003; Boero et al., 2015). However, the strength of the uncertainty-disagreement linkage may differ for more persistent outcome variables such as the unemployment rate. Hence, the analysis from Chapter 4 accounts for differences across outcome variables. Fifth, the theoretical forecast error decomposition of Lahiri and Sheng (2010) predicts that the difference between disagreement and uncertainty increases with the forecast horizon due to an accumulation of aggregate future shocks. Therefore, we evaluate whether uncertainty and disagreement are more closely related at specific forecast horizons.

In line with the evidence from Chapter 3, we find that disagreement is a poor proxy for uncertainty. However, we show that the strength of the correlation varies with the employed dispersion statistic, the usage of either point forecasts or histogram means to calculate disagreement, the considered outcome variable, and the forecast horizon. In contrast, distributional assumptions do not appear to be very influential. Moreover, the relationship is weaker during economically turbulent periods when indicators of uncertainty are needed most. Accounting for the entry and exit of forecasters to and from the survey has little impact on the results.

We also show that survey-based uncertainty is associated with overall policy uncertainty (as measured by EPU), whereas forecaster disagreement is more closely related to the anticipated fluctuations on financial markets (as indicated by VSTOXX). The same factors that are responsible for the decoupling of policy and financial uncertainty may thus also explain the divergence between survey-based measures of uncertainty and disagreement since the outbreak of the Great Recession. In summary, our findings from Chapters 3 and 4 highlight the importance of relying on the subjective probability distributions in order to quantify uncertainty.

Overconfidence Versus Rounding in Survey-Based Density Forecasts

Chapter 5 is concerned with the informative content of the survey predictions for the conditional variance, which has been recently contested, e.g., by Clements (2016). In particular, our analysis contributes to a growing body of literature that analyzes the misalignment of ex-ante and ex-post variances. Although survey-based point predictions have been found to outperform successful forecasting models (e.g., Ang et al., 2007), corresponding variance forecasts are frequently diagnosed as heavily distorted. Forecasters who report inconspicuously low ex-ante variances often produce squared forecast errors that are much larger on average. We ask under which conditions the second moments from the SPF data are relatively well aligned with the variability of the prediction errors. We propose to relate the ex-ante variance of forecasters to the properties of their predictions itself which are observed prior to the outcome. These characteristics are the distinct rounding patterns in individual histogram forecasts, which have been documented in several empirical and experimental studies (Manski and Molinari, 2010; Binder, 2017; Boero et al., 2011; Ruud et al., 2014).

We document the novel stylized fact that the variance misalignment in both the ECB- and FED-SPF is related to the rounding behavior of the survey participants. In particular, the deviations of survey participants' forecast variances prior to and after the outcome can partially be ascribed to the response pattern of a large group of forecasters that submit rounded histogram predictions. Discarding responses which are strongly rounded provides an easily implementable correction that (i) can be carried out in real time, i.e., before outcomes are observed, and (ii) delivers a significantly improved match between ex-ante and ex-post forecast variances. According to our estimates, uncertainty about inflation, output growth and unemployment in the U.S. and the Euro area is higher after correcting for the rounding effect. The observed increase in the share of non-rounded responses in recent years also helps to further understand the trajectory of survey-based average uncertainty during the years after the financial and sovereign debt crisis. Our findings are in line with assertions from the previous literature regarding the connection between survey respondents' rounding behavior and their uncertainty about future macroeconomic outcomes. Moreover, we discuss potential explanations for the decision of survey participants whether or not to round their predictions. In particular, we argue that there may exist a close connection between rounding and the importance of judgment relative to the use of (econometric) models in the formation of individual expectations.

Chapter 2

'Déjà Vol' Revisited

.

2.1 Introduction

> "*The puzzle highlighted by the results in this paper is that stock volatility is not more closely related to other measures of economic volatility.*"
>
> (Schwert, 1989, p. 1146)

> "*Because volatility co-varies with business conditions, a tendency exists to suspect that incorporating macroeconomic information should greatly improve longer horizon volatility forecasts. The relatively comprehensive analysis in this paper shows that only modest forecasting gains are possible.*"
>
> (Paye, 2012, p. 545)

The notion that macroeconomic variables should have predictive power for financial volatility has a long tradition in economics. Theoretically, a counter-cyclical behavior of stock market volatility can be rationalized based on the present value models of Campbell (1991) and Campbell and Shiller (1988) but also within the context of more recent models of asset valuation such as Bansal and Yaron (2004), Mele (2007), Bollerslev et al. (2009), David and Veronesi (2013) or Campbell et al. (2018), among others. Against this background, it is astonishing that the empirical evidence for such a link is rather scant. One of the early empirical studies is Schwert (1989). The above quote from Schwert (1989) refers to the puzzling finding that macroeconomic variables appear to be only

This chapter is based on the working paper "'Déjà vol' revisited: Survey forecasts of macroeconomic variables predict volatility in the cross-section of industry portfolios" that I wrote jointly with my supervisor, Prof. Dr. Christian Conrad (cf. Conrad and Glas, 2018).

weakly related to stock market volatility when conditioning on past volatility. Although subsequent studies have frequently revisited the question of whether macroeconomic variables help to forecast volatility, the quote from Paye (2012) illustrates that the empirical evidence is still somewhat at odds with modern approaches to asset valuation. Because volatility forecasts are of great importance for portfolio choice, risk management and the surveillance of risks to financial stability, we take a fresh look at this question.

In contrast to the previous literature which focuses almost exclusively on the realizations of macroeconomic variables, we argue that survey forecasts of these variables are usually more informative about future stock volatility and, what is more, are available in real-time. Using expectations data from the Survey of Professional Forecasters (SPF) of the Federal Reserve Bank of Philadelphia, we contribute to the literature by providing strong evidence that macroeconomic variables do successfully predict quarterly stock volatility in a cross-section of 49 industry portfolios. We consider an AR(2) specification for quarterly realized industry volatility as a benchmark model and test whether additionally including macroeconomic variables improves forecast performance in- and out-of-sample. Our econometric approach follows Paye (2012), which ensures comparability of our results with his. Our findings can be summarized as follows.

First, based on full sample estimates, we show that one-quarter-ahead SPF forecasts of gross domestic product (GDP) and industrial production (IP) growth are significant predictors of future volatility in 24/33 out of the 49 industries. The signs of the estimated effects are in line with a counter-cyclical behavior of volatility, which is consistent with both a higher volatility of dividend news during contractions than during expansions (see, for example, Bansal and Yaron, 2004) and/or more volatile discount rates during bad times (see Mele, 2007). In sharp contrast, realizations of GDP/IP growth are essentially insignificant predictors in all industries. In addition, it is important to notice that the release schedule of quarterly macroeconomic figures rules out the estimation of predictive regressions with realizations in real-time.

Second, we argue that predictive regressions with realized volatility as the dependent variable suffer from measurement error. As discussed in Engle et al. (2013) and Conrad and Schienle (2018), realized volatility should be considered a noisy proxy for the latent long-term component of volatility that is driven by macroeconomic conditions. The measurement error problem suggests that macroeconomic variables may often falsely appear to be insignificant in predictive regressions despite a potentially existing effect. We use the procedure proposed in Conrad and Schienle (2018) in order to test for an existing relationship. This test simply consists of running the same predictive regression as before but with realized volatility being replaced by ‘volatility-adjusted realized volatility’.

The latter is a more precise proxy for the latent long-term volatility which mitigates the measurement error problem. The test results strengthen the evidence in favor of the predictive ability of the macroeconomic variables. We find that forecasts of GDP/IP growth are significant in 38/37 out of the 49 industries. Moreover, the realizations of the macroeconomic variables gain significance as well. For example, in 26 cases the realization of a higher unemployment rate predicts a significant increase in stock volatility.

Third, even when controlling for standard predictors of returns and volatility such as the dividend-price ratio, the price-earnings ratio, the term spread or news implied volatility, the macroeconomic forecasts retain their significance in the predictive regressions. That is, macroeconomic variables contain information beyond that contained in those predictors and past volatility. In line with David and Veronesi (2013), our results suggest that the relation between inflation and stock volatility depends on the prevailing macroeconomic regime. For example, if higher inflation expectations go along with a lower price-earnings ratio, i.e., higher earnings expectations, stock volatility decreases. On the other hand, if inflation expectations are high and at the same time the term spread is low, investors may fear a stagflation regime and, hence, higher inflation expectations are associated with higher stock volatility.

Fourth, motivated by these findings, we run predictive regressions for subsamples of 15 years each. The predictive power of GDP/IP growth for future volatility is strongest for the turbulent 1970s and early 1980s, decreases during the Great Moderation period but upraises again for subsamples that include the Great Recession. As before, we observe a time-varying effect for inflation. While higher inflation rates predominantly go along with higher stock volatility during the stagflation period of the 1970s and during the Volcker disinflation, increasing inflation rates are associated with lower levels of stock volatility before the financial crisis of 2007/8. Finally, we observe an interesting dichotomy for the unemployment rate. While higher realized unemployment rates go along with elevated levels of stock market volatility, higher expected unemployment rates predict lower volatility in recent subsamples. The former effect suggests that higher realized unemployment rates are perceived as bad news for stocks (via the cash flow effect), while the latter effect appears to be driven by monetary policy: The expectation of higher unemployment rates reduces the probability of a monetary tightening and, hence, is good news for stocks (via the discount effect).

Fifth, we evaluate the out-of-sample predictive ability of the macroeconomic variables by means of Giacomini and White (2006) tests. We test the null hypothesis of equal forecast performance with the AR(2) benchmark model. At the 10% significance level, we find that the null hypothesis is rejected for 14/19 out of the 49 industries when including forecasts of GDP/IP growth and for 20/25 industries when including forecasts of inflation or the unemployment rate. Interestingly, we find that macroeconomic expectations are most informative during the onset of recessions and, in particular, during the period after the Great Recession.

In summary, we provide much more optimistic results regarding the predictive power of macroeconomic variables for stock volatility than the previous literature. Our findings are driven by three main insights. First, while the previous literature mainly focuses on realizations of macroeconomic variables, we show that forecasts are usually more informative. Second, we consider the cross-section of industry portfolios instead of the broad stock market. Our results highlight that not all industries are alike, i.e., stock volatility in some industries is closely related to macroeconomic conditions while the volatility in others is not. Averaging over all industries surely makes it more difficult to detect a robust relationship. Third, for GDP/IP growth the strength of the predictive ability varies over time as does the sign of the estimated effect of inflation. What is more, for the unemployment rate realizations and forecasts even have opposite effects.

The remainder of this chapter is organized as follows. Section 2.2 reviews the related empirical literature and Section 2.3 presents some theoretical thoughts on the relation between macroeconomic conditions and stock volatility. We present the econometric framework in Section 2.4 and the data in Section 2.5. The empirical results are discussed in Section 2.6.Section 2.7 provides extensions and robustness checks. Section 2.8 summarizes and concludes.

2.2 Related Literature

In this section, we review the empirical literature on macroeconomic predictors of stock market volatility. Officer (1973) and Schwert (1989) were among the first to investigate this link. More recent papers are Campbell and Diebold (2009), Paye (2012) and Christiansen et al. (2012). The main econometric workhorse in all those papers is a predictive regression with a measure of realized stock volatility as the dependent variable. The question of interest is whether macroeconomic conditions have forecasting power for future stock volatility when controlling for past volatility. While Paye (2012) and Christiansen et al. (2012) find that financial variables such as default spreads and dividend yields have some predictive power for stock volatility, macroeconomic variables such as GDP or IP growth are not found to be useful. Campbell et al. (2001) consider industry- and firm-level volatility and find no evidence that GDP growth predicts those volatilities. Similarly, Chong and Lin (2017) analyze industry-level stock volatility and find that industrial production growth is not a meaningful predictor. More promising results are presented in Campbell and Diebold (2009) based on six-month growth forecasts from the Livingston survey. They show that higher growth expectations forecast lower levels of CRSP-based realized volatility in-sample. However, Campbell and Diebold (2009) do not provide out-of-sample evidence.

A second strand of literature employs GARCH-type component models for volatility. In particular, Engle and Rangel (2008) and Engle et al. (2013) decompose volatility into a short- and a long-term component and provide evidence that low GDP/IP growth and high inflation predict high long-term volatility. Further evidence is provided in Asgharian et al. (2013) and Conrad and Loch (2015). In particular, Conrad and Loch (2015) show that GDP/IP growth and the unemployment rate are lagging with respect to the long-term component of stock volatility and that the SPF expectations contain useful information that is not included in past realizations. Whereas Conrad and Loch (2015) focus on aggregate stock market volatility, we primarily focus on industry portfolio volatility. This will help us understand the cross-sectional heterogeneity in volatility.

At first, it might be puzzling why the evidence based on predictive regressions is much less optimistic than the evidence based on GARCH-type models. Engle et al. (2013) and Conrad and Schienle (2018) explain that this might be due to a measurement error problem in the predictive regressions. Conrad and Schienle (2018) propose a test for checking whether macroeconomic conditions are related to the long-term component of volatility. We make use of their test and show that standard predictive regressions indeed tend to overlook existing relationships.

While we focus on the effects of macroeconomic conditions on financial volatility, other studies investigate the relation between macroeconomic uncertainty and financial volatility or, reversely, the effects of financial volatility on the macroeconomy. For a recent review of the literature see Andersen et al. (2013).

2.3 The Economics of Volatility

Although the evidence in favor of the predictive ability of macroeconomic variables is rather weak, there is strong empirical evidence suggesting that financial volatility behaves counter-cyclical. The present value models of Campbell and Shiller (1988) and Campbell (1991) provide a theoretical framework for intuitively illustrating the mechanics of the relationship between macroeconomic conditions and stock volatility. In our empirical analysis, we will focus on the volatility of industry portfolios. Let $r_{i,t}$ denote the log return on industry portfolio i in period t. Unexpected returns due to news can be decomposed into two surprise components:

$$r_{i,t+1} - \mathbf{E}_t[r_{i,t+1}] \quad = \quad \varepsilon^{div}_{i,t+1} - \varepsilon^{ret}_{i,t+1}, \tag{2.1}$$

where $\varepsilon^{div}_{i,t+1}$ represents revisions in expectations about future dividend payments and $\varepsilon^{ret}_{i,t+1}$ denotes revisions in expectations about future returns. We refer to the first term as the 'cash flow effect' of news and to the second term as the 'discount rate effect' of news. Because the same piece of news can have a positive cash flow effect (i.e., $\varepsilon^{div}_{i,t+1} > 0$) but a negative discount rate effect (i.e., $-\varepsilon^{ret}_{i,t+1} < 0$), the overall effect of news on unexpected returns is often ambiguous ex-ante and may even vary over the business cycle. For example, Boyd et al. (2005) show that bad news about the unemployment rate are typically good news for stocks during expansions but bad news during contractions. During expansions the discount effect dominates: An increasing unemployment rate reduces interest rate expectations and, hence, is good news for stocks. On the other hand, during contractions the cash flow effect dominates: A higher unemployment rate lowers expected future dividend payments and, hence, is bad news for stocks.

The conditional variance of returns can be written as

$$\mathbf{Var}_t[r_{i,t+1}] \quad = \quad \mathbf{Var}_t[\varepsilon^{div}_{i,t+1}] + \mathbf{Var}_t[\varepsilon^{ret}_{i,t+1}] - 2\mathbf{Cov}_t(\varepsilon^{div}_{i,t+1}, \varepsilon^{ret}_{i,t+1}). \tag{2.2}$$

If expected returns are constant, the conditional variance of returns is time-varying only because of time-variation in $\mathbf{Var}_t[\varepsilon^{div}_{i,t+1}]$. In this case, return volatility is counter-cyclical if the volatility of dividend news is counter-cyclical. If expected returns are time-varying, the conditional volatility of news about expected returns, $\mathbf{Var}_t[\varepsilon^{ret}_{i,t+1}]$, can also generate counter-cyclical stock volatility. This will be the case if changes in expected returns are more pronounced during contractions than during expansions. For examples of models that can generate the former or latter effect see Bansal and Yaron (2004) or Mele (2007). In the general equilibrium model of David and Veronesi (2013) changes in stock market volatility are driven by changes in market participants' beliefs about the prevailing economic regime. Agents learn about the current regime by observing inflation, the price-earnings ratio and other variables. In particular, David and Veronesi (2013, p. 687) emphasize the "time-varying signalling role of inflation".

The previous discussion not only suggests a counter-cyclical behavior of return volatility but also a negative relation between returns and volatility. If the conditional volatility of returns goes up/down in contractions/expansions and contractions/expansions are predominantly associated with bad/good news, returns and volatility are negatively related. This observation is often referred to as the volatility-feedback effect (see, for example, Campbell and Hentschel, 1992).

Another implication is that changes in expected macroeconomic conditions should be associated with variation in $\mathbf{Var}_t[\varepsilon^{div}_{i,t+1}]$ and/or $\mathbf{Var}_t[\varepsilon^{ret}_{i,t+1}]$ and, hence, should be able to predict future stock volatility. Below, we briefly discuss what kind of relationship we expect for the four macroeconomic variables that are employed in the empirical analysis.

GDP and IP growth expectations: If and to the extent to which the conditional volatility of cash flow news of a specific industry behaves counter-cyclical, forecasts of higher GDP growth rates should be associated with lower levels of volatility in the respective industry. Nevertheless, the relationship might not be stable over time. For example, the subdued volatility of GDP/IP growth rates during the Great Moderation might have weakened the link between the business cycle and financial volatility. Moreover, if monetary policy is rules-based and responsive to changes in inflation and GDP growth, higher growth expectations go along with the expectation of a tighter monetary policy. If the economy is in an expansion the discount rate effect might dominate, stock prices fall and volatility increases (see Andersen et al., 2007).

Inflation expectations: Compared to the measures of economic activity, the relationship between stock volatility and inflation expectations is more complex. As highlighted by David and Veronesi (2013) higher inflation expectations can be associated with higher or lower stock volatility. For example, David and Veronesi (2013, p. 684) argue that during the early 1980s “investors faced large uncertainty about whether the United States would enter a persistent stagflation regime”. During this period, higher inflation expectations went along with lower growth expectations, declining stock prices and higher stock volatility. This type of reasoning is consistent with Fama and Schwert (1977) who provide evidence for a negative relationship between inflation and stock returns. Fama (1981) argues that this negative relation can be explained by a negative relation between inflation and growth, whereby growth is the fundamental factor that relates to stock prices. During the 1970s and early 1980s, high inflation simply ‘proxies’ low growth. On the other hand, the model of David and Veronesi (2013) predicts that higher inflation can lead to decreasing stock volatility when the market fears a deflationary regime. In this situation higher inflation expectations are signalling that market participants believe that a deflationary regime can be avoided. As a result, (stock prices increase and) stock volatility declines.

The relation between stock volatility and inflation is also likely to depend on the market participants’ beliefs about the central bank’s reaction function. If the central bank follows an inflation objective, then higher inflation expectations should be accompanied by the expectation of higher policy rates in the future (see Engel and West, 2006, Conrad and Lamla, 2010, and Dräger et al., 2016, for theoretical and empirical evidence). In response, stock prices of interest rate sensitive industries might decline and higher inflation expectations can be associated with higher volatility. Such a relationship might be expected, for example, for the years of the Volcker disinflation. In contrast, if monetary policy is less responsive to inflation, and output and inflation are positively related, higher inflation expectations may predict lower financial volatility.

Unemployment rate: Because the unemployment rate is inversely related to the business cycle, we should expect that a higher unemployment goes along with higher volatility in cash flow news. Thus, the unemployment rate should be positively associated with the volatility of cyclical industries. However, according to Boyd et al. (2005) news about higher unemployment rates during expansions are typically good news for the stock market. If monetary policy is forward looking, an increase in the expected unemployment rate will create the expectation of lower policy rates in the future or will reduce the probability of a tightening during an expansion. Hence, we conjecture that an increase in the expected unemployment rate can also be associated with lower volatility.

2.4 Predictive Regressions for Financial Volatility

Our main workhorse for the empirical analysis is the predictive regression. We employ the same specification as in Paye (2012) and model the volatility in industry $i = 1, \ldots, N$ as

$$\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i} \ln(Vol_{i,t}) + \phi_{2,i} \ln(Vol_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}, \tag{2.3}$$

where $Vol_{i,t}$ denotes a certain measure of stock volatility.[1] The predictor X_t is either the realization, x_t, of one of the previously discussed macroeconomic variables, the corresponding nowcast, $x_{t|t}$, or a one-step-ahead forecast, $x_{t+1|t}$, from the Survey of Professional Forecasters. Since the macroeconomic variables are observed at the quarterly frequency, the index $t = 1, \ldots, T$ refers to subsequent quarters. Finally, the error term, $\nu_{i,t+1}$, captures all factors that affect volatility but not included in the model. We are mainly interested in testing the hypothesis $H_0 : \theta_i = 0$ against $H_1 : \theta_i \neq 0$. Thus, our benchmark model is a simple AR(2). Inference is based on autocorrelation- and heteroskedasticity-robust standard errors (Newey and West, 1987). Prior to estimation, we standardize the explanatory variable X_t by dividing it by its standard deviation. To improve the readability, the coefficients and standard errors reported in the tables below are the estimated ones times 100.

[1]The AR(2) specification can be justified based on the use of information criteria for lag length selection. However, all our results are robust to reasonable modifications of the lag length.

The most commonly used measure of the unobservable volatility is *realized volatility*. We allow for a constant, non-zero industry-specific expected return, μ_i, and model industry returns as

$$r_{i,d,t} = \mu_i + \varepsilon_{i,d,t}, \tag{2.4}$$

where $r_{i,d,t}$ denotes the daily log excess return in industry i on trading day $d = 1, \ldots, D_t$ of quarter t. We compute the demeaned excess return in each sector as the residual from the estimate of Eqn. (2.4), i.e., $\hat{\varepsilon}_{i,d,t} = r_{i,d,t} - \hat{\mu}_i$. Realized volatility in industry i and quarter t is defined as the square root of the sum over the squared residuals, i.e.,

$$RV_{i,t} = \sqrt{\sum_{d=1}^{D_t} \hat{\varepsilon}^2_{i,d,t}}. \tag{2.5}$$

Later on, we will consider a measure of idiosyncratic volatility which is based on an extended specification of Eqn. (2.4) that additionally includes the log excess return for the market portfolio,

$$r_{i,d,t} = \mu_i + \beta_i r_{m,d,t} + \eta_{i,d,t}. \tag{2.6}$$

The coefficient of determination from Eqn. (2.6) captures the systematic risk of industry i, which we will refer to as SR_i in the following sections.

2.5 Data

In this section, we describe the return and macroeconomic data that are used in the main part of this chapter. Data that are exclusively employed in Section 2.7 on robustness will be introduced in the respective subsections.

Stock Market and Industry Portfolios

Financial return data are obtained from the Fama-French Data Library. We use daily excess returns of the aggregate stock market (*mkt*) and the daily value-weighted excess returns from the 5 and 49 industry portfolios. Broadly speaking, the market return is the value-weighted excess return of all firms listed in the NYSE, AMEX, or NASDAQ. The 5 industry portfolios are defined as follows: Business equipment (*hitec*), consumer durables (*cnsmr*), healthcare (*hlth*), manufacturing (*manuf*) and other (*other*). A description of the 49 industry portfolios is provided in Table 2.6 in the Appendix.[2] Our sample covers the 1968Q4 to 2017Q4 period, such that $T = 197$. As will be discussed below, the starting point of our sample period is determined by the availability of the macroeconomic expectations data. As before, we denote the log excess returns on day d of quarter t for industry portfolio i by $r_{i,d,t}$ and for the market portfolio by $r_{m,d,t}$. For the market as well as for each industry portfolio, we calculate quarterly realized volatility, $RV_{m,t}$ and $RV_{i,t}$, as described in Eqns. (2.4) and (2.5). Table 2.1 provides summary statistics for the market and the 5 industry portfolios. The portfolios are sorted in decreasing order according to their systematic risk, SR_i. We retain this sorting throughout the following sections. Table 2.1 presents the systematic risk as well as the corresponding estimate of β_i, the average annualized return and the average annualized realized volatility of each industry. The other measures that are presented in the table will be introduced and discussed later.

Table 2.1: Summary statistics for the market and the 5 industry portfolios

	SR_i	$\hat{\beta}_i$	$\bar{r}_i$	$\overline{RV}_i$	$\overline{\widetilde{RV}}_i$	$\overline{IVol}_i$	$\overline{\widetilde{IVol}}_i$
mkt	1.00	1.00	4.90	14.34	14.25	–	–
other	0.87	1.07	4.29	15.86	15.82	5.58	6.05
cnsmr	0.85	0.87	5.92	13.93	14.34	5.19	5.46
manuf	0.84	0.91	5.31	14.13	14.56	5.31	7.06
hitec	0.83	1.13	4.40	17.70	18.02	7.23	7.43
hlth	0.67	0.88	6.40	16.09	16.57	9.06	9.82

Notes: This table displays the systematic risk (SR_i) and the estimate of β_i from the model in Eqn. (2.6) for the market (*mkt*) as well as the other (*other*), consumer durables (*cnsmr*), manufacturing (*manuf*), business equipment (*hitec*) and healthcare (*hlth*) industries. Moreover, the table contains the annualized time series averages of the quarterly excess return, the realized volatilities ($RV_{i,t}$, $\widetilde{RV}_{i,t}$), and the idiosyncratic volatilities ($IVol_{i,t}$, $\widetilde{IVol}_{i,t}$). The sample period is 1968Q4–2017Q4.

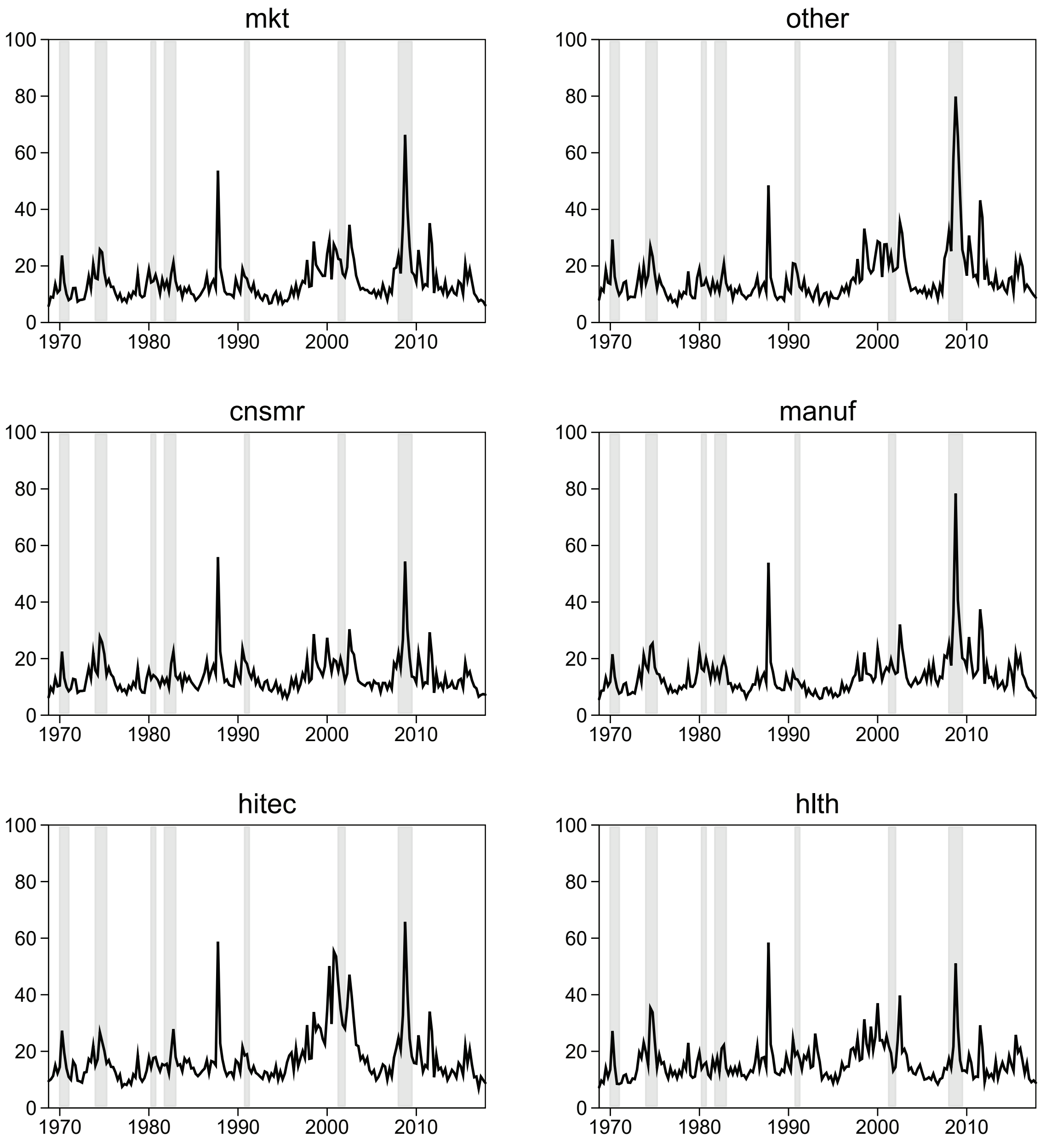
Figure 2.1 depicts the annualized time series of quarterly realized volatility.[3] While the first panel shows the market, the other panels show the 5 industry portfolios. Shaded gray areas indicate recession periods as classified by the National Bureau of Economic Research (NBER).

As expected, all time series of realized volatility are characterized by a distinct countercyclical behavior, i.e., volatility is high during economic contractions and low during expansion periods. While there is an obvious comovement in the realized volatilities of

[2]A more detailed description/definition of 'the market' and the industry portfolios is available on the website of the Fama-French Data Library.

[3]The time series of quarterly returns are depicted in Figure 2.22 in the Appendix.

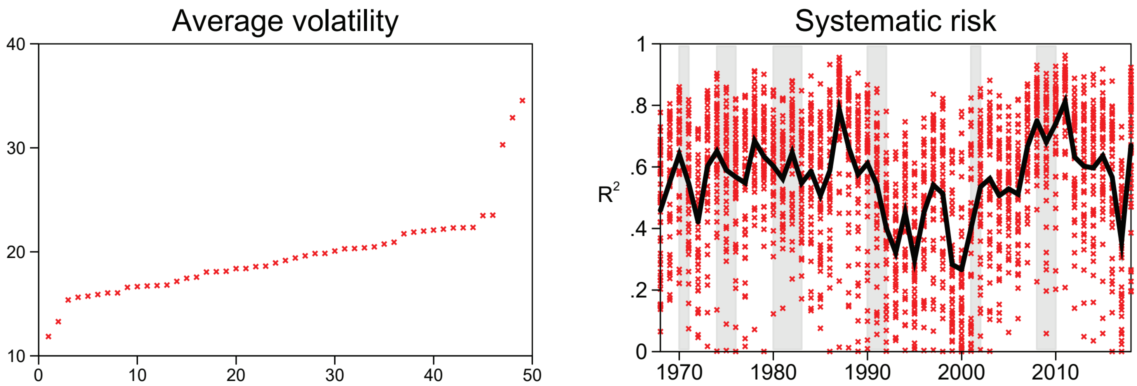
Figure 2.1: Realized volatility for the market and the 5 industry portfolios



Notes: The plots depict the annualized time series of quarterly realized volatility, i.e., $\sqrt{4} \cdot RV_{i,t}$, for the market (*mkt*) as well as the other (*other*), consumer durables (*cnsmr*), manufacturing (*manuf*), business equipment (*hitec*) and healthcare (*hlth*) industries. Sectors are listed in decreasing order according to the systematic risk from Eqn. (2.6) (see Table 2.1). The data are taken from the Fama-French data library. The sample period is 1968Q4–2017Q4. Shaded gray areas indicate NBER-based recession periods.

the 5 industry portfolios, there are also clear idiosyncracies such as the marked increase in volatility in the *business equipment* industry during the period of the dot-com bubble and its final burst. We illustrate the cross-sectional heterogeneity in the quarterly realized volatilities of the 49 industry portfolios in Figure 2.2.

Figure 2.2: Average realized volatility and systematic risk for the 49 industry portfolios



Notes: The left panel depicts the annualized time series average of quarterly realized volatility, i.e., $\overline{RV}_i = (1/T)\sum_{t=1}^{T} RV_{i,t}$ (vertical axis), for each sector from the 49 industry portfolios (horizontal axis), ordered by magnitude. For each year y, the right panel depicts the cross-section of systematic risk $SR^2_{i,y}$ from the regression $r_{i,d,y} = \mu_{i,y} + \beta_{i,y} r_{m,d,y} + \eta_{i,d,y}$ using daily data for each of the 49 industry portfolios (red crosses). The solid black line is the cross-sectional average systematic risk, i.e., $\overline{SR}_y = (1/49)\sum_{i=1}^{49} SR_{i,y}$. The data are taken from the Fama-French data library. The sample period is 1968Q4–2017Q4. Shaded gray areas indicate NBER-based recession periods. Year y is considered to be a recession year if at least one of the four quarters is defined as a recession period by the NBER.

The left panel shows that the average annualized $RV_{i,t}$ is around 12% for 'low-volatility industries' and above 30% for 'high-volatility industries'.[4] For comparison, the average annualized realized volatility of the market is 14.34% (see Table 2.1). The right panel is based on daily regressions of industry excess returns on a constant and the market's excess return. Within each year, we estimate the systematic risk of all 49 industries using daily data and then average over the industries. The graph shows the individual systematic risks for the 49 industries (red crosses) and the evolution of the cross-sectional average systematic risk (black line) over time. For a sample that ends in 1997, Campbell et al. (2001) provide a similar figure and find that average systematic risk is trending downwards (see their Figure 5). The right panel of Figure 2.2 shows that this was only a temporary phenomenon and that the average systematic risk has spiked again around 2010. We conjecture that the ability of macroeconomic variables to forecast market volatility may be positively related to the average systematic risk and, hence, may vary over time. We

[4]The lowest/highest average annualized realized volatility is observed in the *utilities* (11.86%) and *precious metals* (34.55%) industries, respectively.

investigate this issue in Section 2.6.4.

Macroeconomic Expectations Data

Data on survey expectations and realizations of the macroeconomic variables are taken from the Federal Reserve's SPF. We focus on four key variables that are available on a quarterly basis from 1968Q4 onwards. We consider two measures of economic activity: Real GDP growth (Δgdp) and industrial production growth (Δip). Growth rates are defined as annualized quarter-over-quarter percentage changes. Inflationary developments are proxied by percentage changes in the GDP price index (inf). Finally, we employ the change (first difference) of the civilian unemployment rate (Δune).

For each of the four variables, we compute the growth rates/changes based on the median of the predictions reported by the individual survey respondents. The SPF is conducted in the middle month of each quarter and consists of approximately 35–45 participants per questionnaire. Our analysis is based on the nowcasts for the current quarter, $x_{t|t}$, as well as on the one-quarter-ahead forecasts, $x_{t+1|t}$.

In addition to the expectations, the SPF provides different data vintages of the realizations for each outcome variable, x_t. We employ first-release data of the realizations.[5] It is important to note that, for example, the first release of GDP growth in quarter t becomes available in quarter $t+1$. Thus, strictly speaking the predictive regression in Eqn. (2.3) is infeasible in real-time when using the realizations x_t.[6] In contrast, predictive regressions based on $x_{t|t}$ and $x_{t+1|t}$ are feasible in real-time.

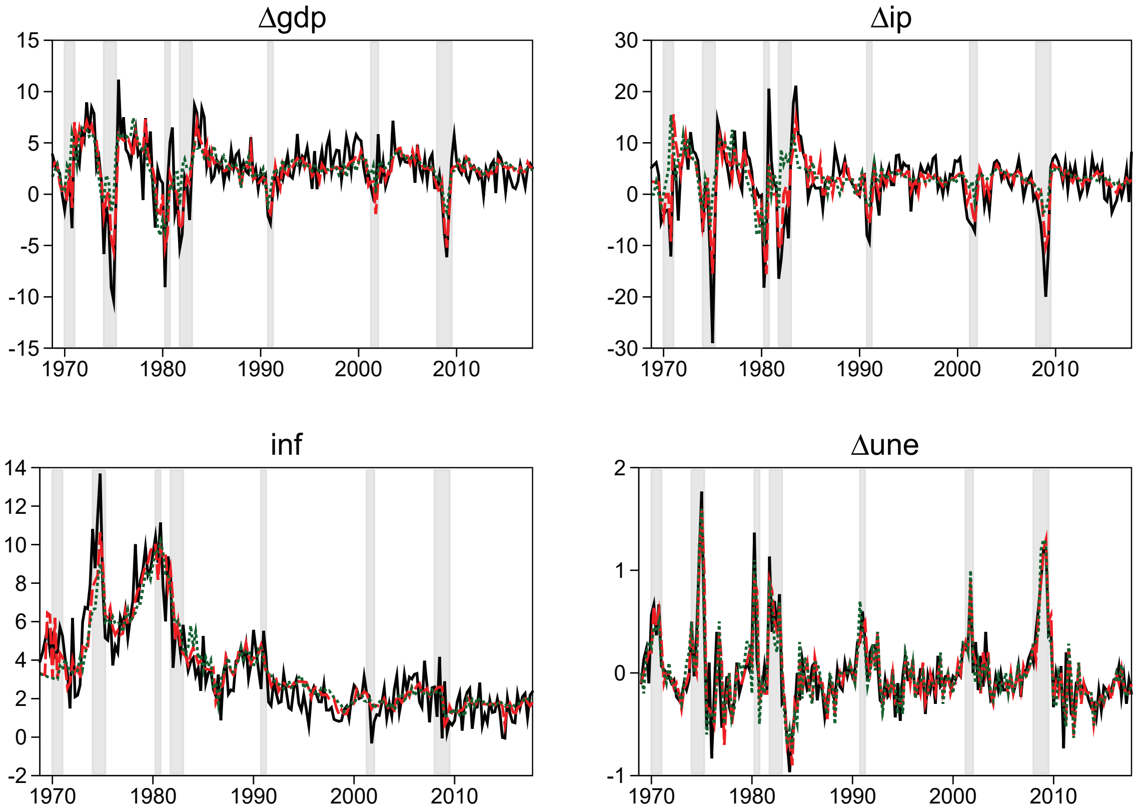
Figure 2.3 shows the realizations (black solid line), nowcasts (red dashed line) and forecasts (green dotted line) of the macroeconomic variables. The horizontal axis depicts the period during which the predictions are reported and realizations are observed. All four variables are volatile in the 1970s and early 1980s, but become increasingly tranquil during the Great Moderation period which ends with the Great Recession.

Table 2.2 provides summary statistics. For all variables but the unemployment rate, the mean of the one-quarter-ahead forecasts is above the mean of the realizations. That is, on average survey participants are too optimistic with respect to GDP/IP growth. At the same time, the standard deviation of the forecasts is lower than the standard deviation of the realizations. In addition, forecasts have a lower interquartile range and a lower kurtosis than the realizations. Finally, forecasts are more persistent than realizations.

[5]We replace the missing observations for the first releases of Δgdp and inf in 1995Q4 with the corresponding values from the second releases.

[6]Note that this problem is less severe for industrial production and the unemployment rate where the quarterly observations are based on data that is released on a monthly basis. For example, at the end of quarter t, industrial production figures for the first two months of quarter t are already available and only the figure for the last month is missing.

Figure 2.3: Forecasts and realizations of macroeconomic variables



Notes: Plots depict the time series of the realizations (x_t, solid black line) and the annualized percentage change of the median nowcasts ($x_{t|t}$, dashed red line) and one-quarter-ahead median forecasts ($x_{t+1|t}$, green dotted line) for real GDP (Δgdp), industrial production (Δip) and inflation (inf) as well as the change in the median forecast of the civilian unemployment rate (Δune). The horizontal axis indicates the period during which predictions are reported and realizations are observed. The data is taken from the Federal Reserve's SPF. The sample period is 1968Q4–2017Q4. Shaded gray areas indicate NBER-based recession periods.

Additional Predictor Variables

In addition to the macroeconomic variables, we consider a set of variables that are frequently associated with financial volatility in both the theoretical and empirical literature (see, for example, Christiansen et al., 2012, David and Veronesi, 2013, and Mele, 2007).

We include the (log) dividend-price ratio (dpr) and the cyclically adjusted (log) price-earnings ratio (per). The data are taken from Robert Shiller's website. In addition, we consider the term spread (ts), which is defined as the difference between 10-year government bonds and 1-year T-bill rates. Moreover, we use the News Implied Volatility Index (NVIX, $nvix$) proposed by Manela and Moreira (2017) which measures (investor) uncertainty based on newspaper articles. The NVIX is based on machine learning techniques

Table 2.2: Summary statistics for the macroeconomic explanatory variables

	Variable	Timing	Mean	Std. dev.	Skew.	Kurt.	$q_{0.25}$	$q_{0.75}$	ρ_1
		t	2.39	3.02	−0.96	6.33	1.22	3.90	0.49
Δgdp	Real GDP growth	$t\|t$	2.34	2.19	−0.92	5.14	1.56	3.31	0.71
		$t+1\|t$	2.64	1.76	−0.70	5.44	2.17	3.37	0.80
		t	2.15	6.32	−1.08	7.10	−0.21	5.57	0.53
Δip	Industrial production	$t\|t$	2.45	4.39	−1.06	6.45	1.33	4.37	0.56
		$t+1\|t$	3.25	3.29	−0.05	5.68	2.23	4.26	0.69
		t	3.45	2.59	1.33	4.58	1.63	4.50	0.83
inf	Inflation rate	$t\|t$	3.54	2.26	1.31	3.95	1.80	4.34	0.94
		$t+1\|t$	3.49	2.11	1.29	3.92	1.91	4.30	0.97
		t	0.00	0.37	1.47	7.22	−0.20	0.10	0.51
Δune	Unemployment rate	$t\|t$	0.00	0.36	1.41	6.31	−0.20	0.10	0.61
		$t+1\|t$	0.00	0.36	1.50	6.73	−0.20	0.10	0.55

Notes: This table displays summary statistics for the realizations (x_t), nowcasts ($x_{t|t}$) and one-quarter-ahead forecasts ($x_{t+1|t}$) of the macroeconomic variables from the SPF data. For each variable the columns report the mean, standard deviation, skewness, kurtosis, lower and upper quantiles and the first-order autocorrelation coefficient. The sample period is 1968Q4–2017Q4.

and resembles the behavior of the Volatility Index released by the Chicago Board Options Exchange which is only available from 1990 onwards. Further, we consider the quarterly consumption-wealth ratio (cay) of Lettau and Ludvigson (2001), which is defined as the residual from an estimated cointegration relationship between aggregate consumption, wealth and labor income.[7] Finally, we consider the Index of Consumer Sentiment (ics) from the Michigan Survey of Consumers. Data on cay are available through 2016Q4 and the time series for $nvix$ ends in 2016Q1. All other variables are observed throughout the entire sample period.

2.6 Empirical Analysis

In this section, we discuss the main empirical results. Full-sample predictive regressions for the 5 and 49 industry portfolios are presented in Section 2.6.1. The Conrad and Schienle (2018) test is implemented in Section 2.6.2. In Section 2.6.3, we illustrate that our findings are robust to the inclusion of further predictor variables. Rolling-window estimates and an out-of-sample forecast evaluation are presented in Sections 2.6.4 and 2.6.5, respectively.

[7]We use an updated version of the data used in Goyal and Welch (2008).

2.6.1 Predictive Regressions Based on Realized Volatility

In the first step, we discuss the ordinary least squares (OLS) estimates of the predictive regressions for realized volatility based on Eqn. (2.3).

For illustrative purposes, we begin by discussing the findings for the broad stock market and the 5 industry portfolios. The first double-column in Table 2.3 displays the results for the market, the other double-columns the estimates for the individual industries. In each row, we present the parameter estimate, $\hat{\theta}_i$, based on either (i) the realization, x_t, (ii) the nowcast, $x_{t|t}$, or (iii) the one-quarter-ahead forecast, $x_{t+1|t}$, of the respective macroeconomic variable. Recall that we standardize each variable X_t prior to estimation which ensures comparability of the coefficients. In addition to the parameter estimates, we present the percentage increase in the R^2 (denoted by ΔR^2) relative to the AR(2) benchmark.

In line with the previous literature, the realizations of the macroeconomic variables have no predictive power for future market volatility when controlling for past volatility. As mentioned before, predictive regressions based on x_t are infeasible in real-time. Hence, we replace the realizations with the SPF nowcasts. However, none of the nowcasts significantly affect market volatility. Next, we consider the one-quarter-ahead forecasts of the macroeconomic variables. In this case, we obtain negative and highly significant estimates of θ_m for GDP and IP growth which is in line with the notion of a countercyclical behavior of volatility. The magnitude of the estimated coefficients is also economically significant. For example, a one standard deviation increase in expected real GDP growth is associated with a predicted decline in future market volatility by 4.07% $(100 \cdot [\exp(\hat{\theta}_m \cdot \Delta x_{t+1|t}) - 1] = 100 \cdot [\exp(-0.0416 \cdot 1) - 1] = -4.07\%)$. When including the SPF expectations $x_{t+1|t}$, the percentage change in R^2 is roughly 2% for GDP growth and 2.4% for IP growth.[8] In contrast, neither the inflation nor the unemployment expectations are found to be significant.

Our finding that realizations of GDP growth and IP growth do not forecast stock market volatility are in line with the results in Paye (2012). The evidence that expectations regarding future developments of GDP growth and IP growth are useful predictors for stock market volatility squares with the results in Campbell and Diebold (2009) and Conrad and Loch (2015).[9]

The results for the 5 industry portfolios are in line with those for the market. Again, we find that realizations and nowcasts of the macroeconomic variables are insignificant in

[8] For reference, the goodness of fit for the market portfolio from the AR(2) benchmark is 0.44.

[9] To the contrary, Paye (2012) uses 6- to 12-months GDP growth forecasts from the Livingston Survey and finds no evidence of predictive power of the forecasts.

Table 2.3: Predictive regressions for realized volatility (market and 5 industry portfolios)

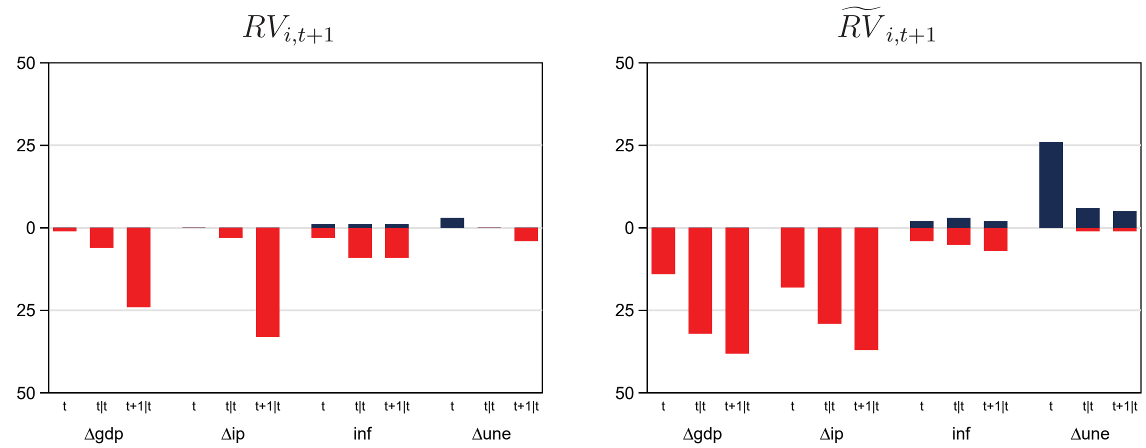
		Industry portfolio											
		mkt		*other*		*cnsmr*		*manuf*		*hitec*		*hlth*	
Predictor x	Timing	$\hat{\theta}_m$	ΔR^2_m	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i
Δgdp	t	−0.45	0.03	−0.32	0.01	−0.73	0.11	−0.77	0.06	0.12	0.00	−0.03	0.00
		(1.47)		(1.20)		(1.12)		(1.70)		(1.66)		(1.91)	
Δgdp	$t\|t$	−2.15	0.51	−2.12	0.39	−2.00	0.74	−2.19	0.47	−0.58	0.03	−1.92	0.80
		(1.72)		(1.50)		(1.25)		(1.80)		(1.75)		(1.64)	
Δgdp	$t+1\|t$	−4.16***	2.06	−3.81**	1.31	−3.56***	2.51	−4.64***	2.23	−2.69	0.73	−2.99**	2.05
		(1.45)		(1.76)		(1.26)		(1.66)		(1.79)		(1.35)	
Δip	t	0.08	0.00	−0.21	0.00	0.43	0.04	−0.55	0.03	0.76	0.06	0.27	0.02
		(2.12)		(1.79)		(1.78)		(1.91)		(1.95)		(2.23)	
Δip	$t\|t$	−1.00	0.12	−0.93	0.08	−0.11	0.00	−1.70	0.31	0.03	0.00	−0.85	0.16
		(1.82)		(1.54)		(1.40)		(1.88)		(1.77)		(1.39)	
Δip	$t+1\|t$	−4.35***	2.39	−4.43***	1.87	−3.43***	2.41	−4.64***	2.42	−3.45**	1.23	−3.48***	2.86
		(1.24)		(1.42)		(1.08)		(1.65)		(1.46)		(1.08)	
inf	t	−0.34	0.02	−2.32	0.55	0.73	0.12	0.58	0.04	−1.29	0.18	−0.32	0.03
		(1.51)		(1.48)		(1.86)		(1.36)		(1.50)		(1.97)	
inf	$t\|t$	−0.63	0.06	−2.40*	0.57	−0.07	0.00	0.05	0.00	−1.14	0.13	−0.28	0.02
		(1.46)		(1.32)		(1.52)		(1.44)		(1.39)		(1.63)	
inf	$t+1\|t$	−0.39	0.02	−2.14*	0.44	0.18	0.01	0.27	0.01	−0.99	0.10	−0.10	0.00
		(1.38)		(1.20)		(1.52)		(1.39)		(1.35)		(1.59)	
Δune	t	2.36	0.65	2.21	0.43	2.57	1.27	2.34	0.58	1.17	0.13	1.71	0.70
		(1.94)		(2.10)		(1.75)		(1.80)		(2.20)		(2.10)	
Δune	$t\|t$	−1.46	0.25	−1.31	0.15	−1.31	0.32	−1.01	0.11	−2.62	0.65	−1.67	0.66
		(2.04)		(1.95)		(1.76)		(2.21)		(1.86)		(2.16)	
Δune	$t+1\|t$	−2.54	0.74	−2.37	0.48	−2.55	1.21	−1.60	0.27	−3.49*	1.16	−2.44	1.38
		(2.05)		(1.99)		(1.75)		(2.15)		(1.95)		(2.02)	

Notes: This table displays the estimated slope coefficients $\hat{\theta}_i$ from the regression $\ln(RV_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(RV_{i,t}) + \phi_{2,i}\ln(RV_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the market portfolio and the 5 industry portfolios, where $RV_{i,t+1}$ denotes realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization, the nowcast or the forecast of the respective macroeconomic variable. The percentage increase in the R^2 relative to the AR(2) benchmark is denoted by ΔR^2. The R^2_m for the market portfolio based on the AR(2) benchmark is 0.44. The constant and the coefficients on the autoregressive terms are not reported. The predictor X_t is standardized with respect to its standard deviation. Industries are listed in decreasing order according to the SR_i from Eqn. (2.6) (see Table 2.1). The estimation sample $t = 1, \ldots, 197$ covers the period 1968Q4–2017Q4. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. The coefficients and standard errors are the estimated ones times 100. Asterisks *, **, and *** indicate significance at the 10%, 5% and 1% critical level, respectively.

essentially all cases. To the contrary, GDP growth expectations for $t+1$ are significantly negative for four and IP growth expectations for all five industry portfolios. Both GDP and IP growth expectations have the strongest effect on the *manufacturing* industry. We also observe notable increases in the R^2 of the predictive regressions when adding $x_{t+1|t}$.[10] For example, for one-quarter-ahead GDP growth forecasts the increase in the R^2 is 2.51% for the *consumer durables* sector and for IP growth forecasts the increase is 2.86% for the *healthcare* industry. Thus, the predictive regressions for the industry portfolios confirm our results for the stock market, but also illustrate that the estimates of θ_i as well as the percentage increases in R^2 vary across industries.

To further investigate the relation between macroeconomic conditions and stock volatility in the cross-section of industries, we now consider the 49 industry portfolios. To simplify the interpretation, the estimation results from the predictive regressions are presented graphically. The left panel of Figure 2.4 depicts the number of estimates of θ_i that are significantly different from zero at the 5% critical level for each macroeconomic variable $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$. The blue/red bars indicate significantly positive/negative estimates.

Figure 2.4: Number of significant estimates of θ_i for the 49 industry portfolios ($RV_{i,t+1}$ and $\widetilde{RV}_{i,t+1}$)



Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the 49 industry portfolios based on predictors $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ when either realized volatility ($RV_{i,t+1}$, left panel) or volatility-adjusted realized volatility ($\widetilde{RV}_{i,t+1}$, right panel) is considered as the volatility proxy. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

In line with the evidence from Table 2.3, we find that future volatility is almost always unrelated to the realizations of GDP and IP growth. The association between realized

[10] The R_i^2 from the AR(2) benchmark range from 0.32 (*healthcare*) to 0.54 (*business equipment*).

volatility, inflation and the unemployment rate appears to be weak as well. The effect of inflation is significantly negative in only three out of the 49 industry portfolios and an increase in the unemployment rate tends to significantly increase future volatility in three industries. When using nowcasts instead of realizations, we observe a few more significant estimates for GDP/IP growth as well as for inflation. In line with the results for the 5 industry portfolios, forecasts of GDP/IP growth are much stronger predictors of realized volatility. GDP/IP growth significantly predict lower volatility in 24/33 industries. Interestingly, we observe that higher unemployment rate forecasts now predict decreasing realized volatility in four industries. We will discuss and explain this switch in signs in detail in Section 2.6.4, which presents rolling-window estimates of the predictive regressions.

The panels in Figure 2.5 depict the percentage improvement in the goodness of fit compared to the baseline AR(2) model when including either the realization (horizontal axis) or the forecast (vertical axis) of the respective macroeconomic variable as a predictor in Eqn. (2.3).[11] The R_i^2 from the AR(2) benchmark range from 0.26 (*medical equipment*) to 0.69 (*computer software*).[12] Red dots indicate that in the underlying regression, θ_i is significantly different from zero at the 5% level either for x_t, $x_{t+1|t}$, or both.

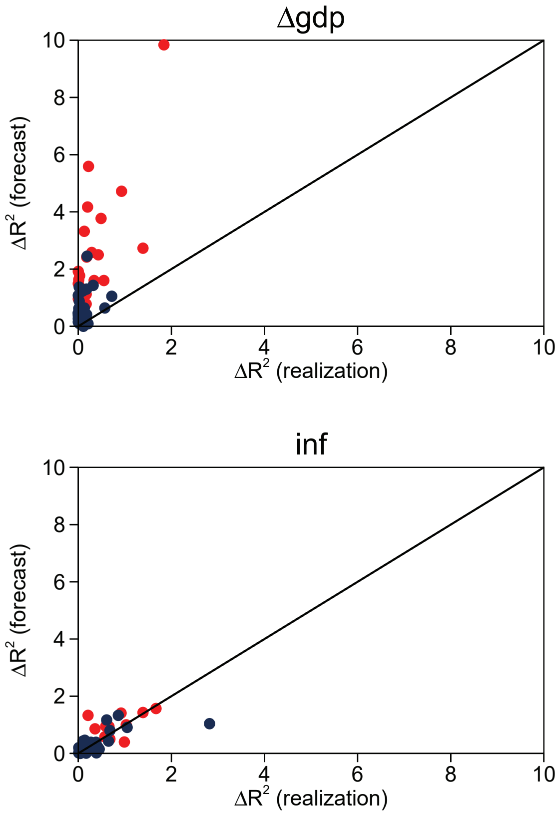
In line with our previous considerations, Figure 2.5 clearly shows that the one-step-ahead SPF forecasts of GDP and IP growth lead to higher percentage increases in R^2 than the corresponding realizations. For both, GDP and IP growth the largest percentage increase in R^2 is observed for the *ships* industry (9.84% and 6.90%, respectively). For inflation, forecasts as well as realizations lead to percentage increases of comparable size. In contrast, the evidence for the unemployment rate is mixed. For some industries, we observe the strongest increases for the realizations (e.g., *medical equipment*: 6.99%) while in other industries for the forecasts (e.g., *retail*: 3.10%).

In sum, our results show that forecasts of GDP and IP growth are strong predictors of future volatility in roughly 50% of the industry portfolios. The size of the estimated effects as well as the percentage change in the goodness-of-fit (as compared to the AR(2) benchmark) suggest that the increase in predictive ability is economically significant. In Sections 2.6.3 and 2.6.4, we provide explanations for why the results for inflation and unemployment forecasts appear to be more modest. In particular, we show that for these two variables the signs of the effects change over time and, hence, full sample regressions may be misleading.

[11] We have also considered the improvement in the goodness of fit for the nowcasts. A detailed account is omitted here for brevity.

[12] The (financial) *trading* sector, which may be of particular interest, has a benchmark R_i^2 of 0.67.

Figure 2.5: Gain in the goodness of fit for the 49 industry portfolios ($RV_{i,t+1}$)



Notes: The plots depict the percentage increase in R_i^2 relative to the AR(2) benchmark for the 49 industry portfolios when either the forecast ($x_{t+1|t}$ vertical axis) or the realization (x_t, horizontal axis) of the respective macroeconomic variable is considered as the predictor in the model from Eqn. (2.3) and realized volatility ($RV_{i,t+1}$) is considered as the volatility proxy. Red dots indicate that in the underlying regression, θ_i is significantly different from zero at the 5% level either for x_t, $x_{t+1|t}$, or both. The estimation sample covers the period 1968Q4–2017Q4.

2.6.2 Volatility-Adjusted Realized Volatility

Based on the representation in Eqn. (2.1), Engle and Rangel (2008) and Engel et al. (2013) suggest that daily unexpected returns can be modeled as

$$\varepsilon_{i,d,t+1}^{div} - \varepsilon_{i,d,t+1}^{ret} = \varepsilon_{i,d,t+1} = \sqrt{\tau_i(\mathbf{s}_{i,t+1})h_{i,d,t+1}}z_{i,d,t+1}, \tag{2.7}$$

where $z_{i,d,t+1}$ represents (industry-specific) news and $\tau_i(\mathbf{s}_{i,t+1})h_{i,d,t+1}$ is the time-varying impact multiplier, which consists of two multiplicative components. The $\tau_i(\mathbf{s}_{i,t+1})$ component depends on low-frequency state variables, $\mathbf{s}_{i,t+1}$, and changes at the quarterly frequency only, while $h_{i,d,t+1}$ represents day-to-day changes in the impact multiplier. Following Engle and Rangel (2008), we refer to $\tau_i(\mathbf{s}_{i,t+1})$ and $h_{i,d,t+1}$ as the long- and short-term volatility component, respectively. Naturally, $\mathbf{s}_{i,t+1}$ depends on variables that proxy the state of the macroeconomy. As discussed in Section 2.3, we might expect that, for example, the same piece of news has a stronger effect in a recession than in an expansion. Alternatively, $\sqrt{h_{i,d,t+1}}z_{i,d,t+1}$ can also be viewed as representing news with time-varying intensity $h_{i,d,t+1}$, where $h_{i,d,t+1}$ follows a GARCH (or some other conditionally heteroskedastic) process.

If Eqn. (2.7) holds, the predictive regression in Eqn. (2.3) with $RV_{i,t+1}$ as the volatility proxy can be interpreted as the linear projection of

$$\ln(RV_{i,t+1}) = \frac{1}{2}\ln\left(\tau_i(\mathbf{s}_{i,t+1})\right) + \frac{1}{2}\ln\left(\sum_{d=1}^{D_{t+1}} h_{i,d,t+1}z_{i,d,t+1}^2\right) \tag{2.8}$$

on a constant and $\mathbf{s}_{i,t+1} = (\ln(RV_{i,t}), \ln(RV_{i,t-1}), X_t)'$. While the first term in Eqn. (2.8) depends on X_t, the second term depends on the sum of a squared daily GARCH process. Thus, as discussed in Engel et al. (2013) and Conrad and Schienle (2018), the dependent variable, $\ln(RV_{i,t+1})$, is a noisy proxy for the variable of interest, $\ln\left(\tau_i(\mathbf{s}_{i,t+1})\right)$. Clearly, the presence of the strongly persistent measurement error, $\ln\left(\sum_{d=1}^{D_{t+1}} h_{i,d,t+1}z_{i,d,t+1}^2\right)$, will tend to mask the existence of a relationship between long-term volatility and X_t.

Conrad and Schienle (2018) propose a test for the existence of a time-varying long-term component that is driven by X_t. In their framework, $\tau_{i,t} = \tau_i(X_t)$, and the null hypothesis is $H_0 : \tau_{i,t} = \tau_i$, where τ_i would be an industry-specific constant. The actual test can be implemented by a two-step procedure. In the first step and under the null hypothesis, we estimate the following GJR-GARCH specification for each industry:

$$\begin{aligned} r_{i,d,t} &= \mu_i + \varepsilon_{i,d,t}, && (2.9)\\ \varepsilon_{i,d,t} &= \sqrt{h_{i,d,t}}z_{i,d,t}, && (2.10)\\ h_{i,d,t} &= \omega_i + (\alpha_i + \gamma_i \cdot \mathbb{1}(\varepsilon_{i,d-1,t} < 0))\varepsilon_{i,d-1,t}^2 + \delta_i h_{i,d-1,t}, && (2.11) \end{aligned}$$

where $z_{i,d,t} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. In this model τ_i is given by $\omega_i/(1-\alpha_i-\delta_i-\gamma_i/2)$. Based on the parameter estimates, we obtain the volatility-adjusted residuals,

$$\hat{\tilde{\varepsilon}}_{i,d,t} = \frac{\hat{\varepsilon}_{i,d,t}}{\sqrt{\hat{h}_{i,d,t}}}, \tag{2.12}$$

and construct the volatility-adjusted realized volatility in sector i and quarter t as

$$\widetilde{RV}_{i,t} = \sqrt{\frac{\hat{\omega}_i}{1-\hat{\alpha}_i-\hat{\delta}_i-\hat{\gamma}_i/2}} \cdot \sqrt{\sum_{d=1}^{D_t}(\hat{\tilde{\varepsilon}}_{i,d,t})^2}, \tag{2.13}$$

where the first term is a scaling factor that ensures that $\widetilde{RV}_{i,t}$ is measured on the same scale as $RV_{i,t}$. In the second step, we estimate the predictive regression given in Eqn. (2.3) but with volatility-adjusted realized volatility instead of realized volatility on the left and right hand side. As before, we test the null hypothesis $H_0 : \theta_i = 0$. As shown in Conrad and Schienle (2018), the new predictive regression is much less prone to measurement error and, hence, more powerful.

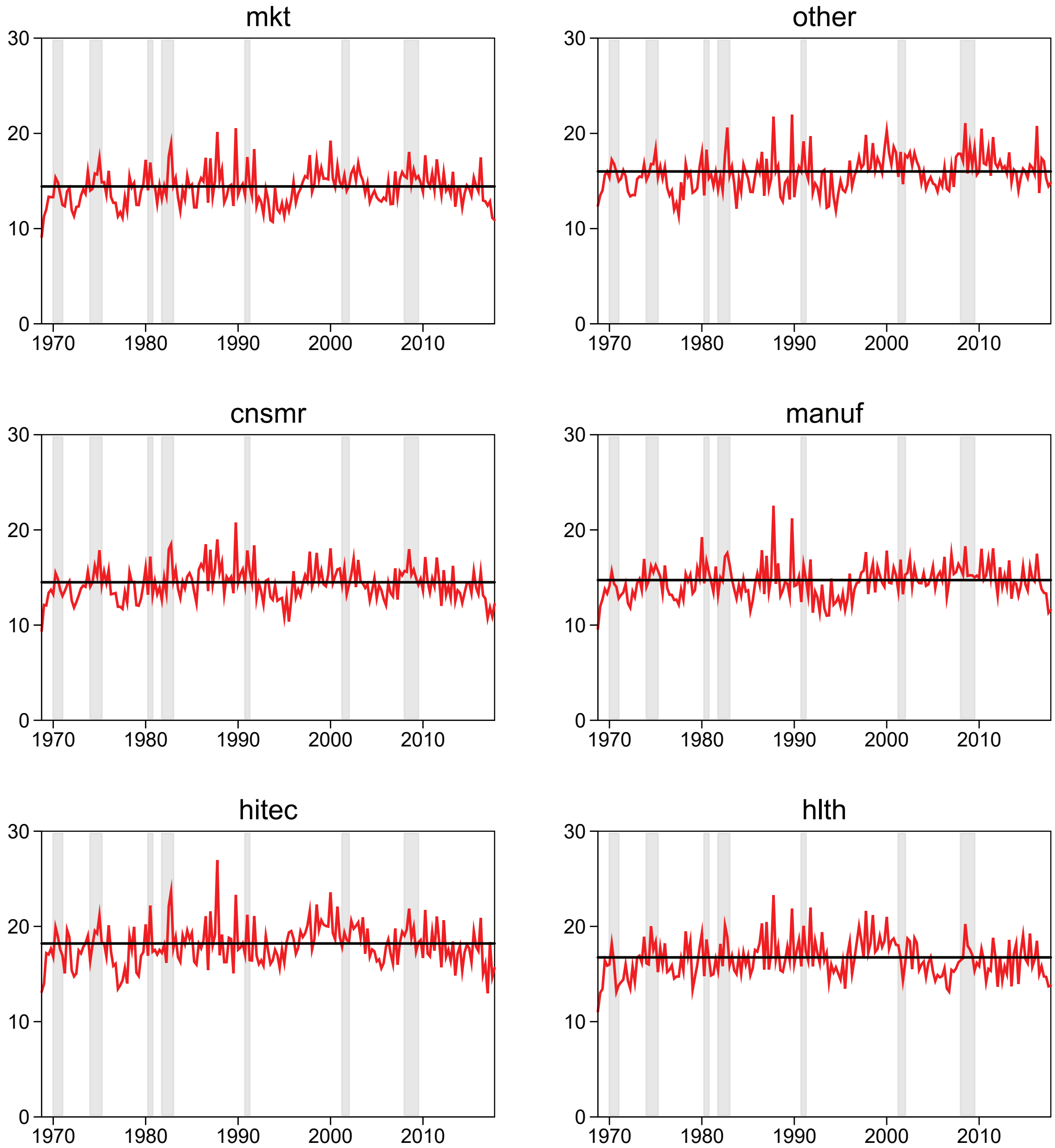
For the market and the 5 industry portfolios, Figure 2.6 shows the annualized time series of the volatility-adjusted realized volatilities, $\widetilde{RV}_{i,t}$ (red lines), along with the annualized estimates of $\sqrt{\tau_i}$ (black lines). While $\widetilde{RV}_{i,t}$ still follows the general trends in realized volatility, it behaves much more smoothly than realized volatility. Table 2.1 shows that for each industry the average of the volatility-adjusted realized volatility is close to the mean of the realized volatility.

The right panel in Figure 2.4 shows the number of significantly positive/negative estimates of θ_i in the 49 industry portfolios when $RV_{i,t+1}$ is replaced with $\widetilde{RV}_{i,t+1}$.[13] The figure shows that for almost all choices of $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$, we now observe more rejections of the null hypothesis than when using $RV_{i,t+1}$ as the volatility proxy (left panel). For example, for expected GDP growth we observe 24 rejections when using $RV_{i,t+1}$ as the volatility proxy whereas we observe 38 rejections in the case of $\widetilde{RV}_{i,t+1}$. Interestingly, when using $\widetilde{RV}_{i,t+1}$, the null hypothesis that x_t does not predict volatility is rejected for 14/18 of the industry portfolios for realized GDP growth, while we hardly observe any rejections in those cases when $RV_{i,t+1}$ is considered. In contrast, for the inflation rate there is no such effect. Finally, when looking at the unemployment rate, we find 26 rejections for x_t in the predictive regression based on $\widetilde{RV}_{i,t+1}$ but only three rejections in the regression based on $RV_{i,t+1}$.

In summary, we find that predictive regressions based on the volatility-adjusted realized volatility lead to much stronger rejections of the null hypothesis that macroeconomic conditions do not affect industry-level volatility. Still, forecasts of GDP/IP growth are most important, while for the unemployment rate the realizations are more relevant. We conclude that there is a severe problem with measurement error in the standard predictive regression with $RV_{i,t+1}$ as the volatility proxy which tends to mask an existing relationship between macroeconomic conditions and volatility. To faciliate the comparison with previous studies, we report the results for both measures of volatility in the remaining subsections.

[13]For completeness, Table 2.7 in the Appendix presents the detailed parameter estimates for the 5 industry portfolios.

Figure 2.6: Volatility-adjusted realized volatility and the 5 industry portfolios



Notes: The plots depict the annualized time series of quarterly volatility-adjusted realized volatility, i.e., $\sqrt{4} \cdot \widetilde{RV}_{i,t}$ (red line), for the market (*mkt*) as well as the other (*other*), consumer durables (*cnsmr*), manufacturing (*manuf*), business equipment (*hitec*) and healthcare (*hlth*) industries. The black lines indicate the annualized estimates of $\sqrt{\tau_i}$. Sectors are listed in decreasing order according to the systematic risk from Eqn. (2.6) (see Table 2.1). The data are taken from the Fama-French data library. The sample period is 1968Q4–2017Q4. Shaded gray areas indicate NBER-based recession periods.

2.6.3 Controlling for Other Predictors

In this section, we investigate whether the predictive power of the macroeconomic variables remains intact when controlling for other variables. That is, we now ask whether the macroeconomic variables contain information about future volatility beyond that contained in variables that have been shown to predict volatility (and/or returns). This question is of particular interest, because the previous literature has found that financial variables (such as the dividend-price ratio) predict volatility while macroeconomic variables "hardly show up as important predictors" (Christiansen et al., 2012, p. 958). We employ an extended specification of the predictive regression in Eqn. (2.3),

$$\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \theta_i X_t + \lambda_i w_t + \xi_{i,t+1}, \qquad (2.14)$$

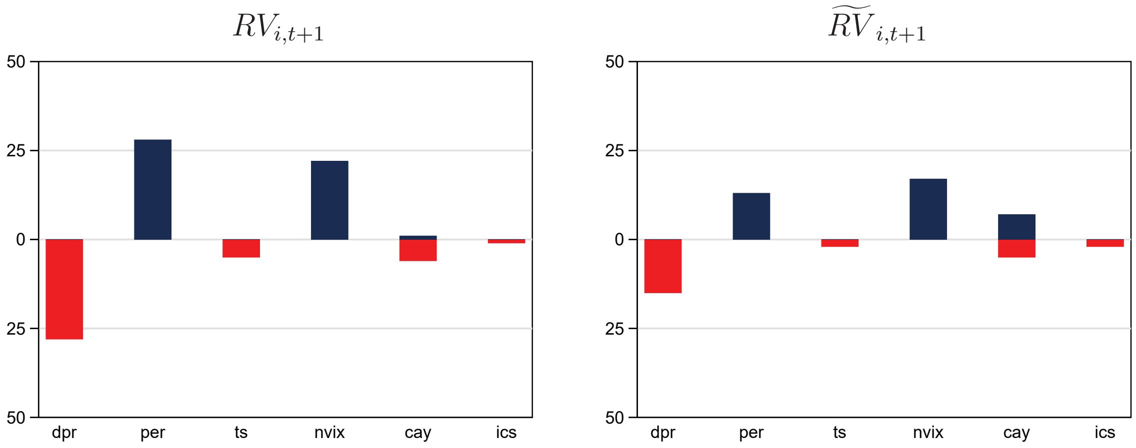
where w_t is one of the following variables: First, we consider financial predictors such as the dividend-price ratio (dpr), the price-earnings ratio (per) and the term spread (ts). Second, we account for the level of uncertainty by including the News Implied Volatility Index ($nvix$). Third, we use the consumption-wealth ratio (cay) and the Index of Consumer Sentiment (ics).

We first check whether the additional variables have predictive power for volatility. To this end, we estimate a version of Eqn. (2.14) that omits the macro variable X_t. Figure 2.7 shows how often these variables significantly predict volatility when either $RV_{i,t+1}$ (left panel) or $\widetilde{RV}_{i,t+1}$ (right panel) is used as the volatility proxy.

Clearly, the dividend-price ratio, the price-earnings ratio as well as the NVIX have the highest predictive ability. The present value model introduced in Section 2.3 implies that a high dividend-price ratio today forecasts higher excess returns in the future.[14] This is consistent with our finding of a negative effect of the dividend-price ratio on future stock volatility. In contrast, a high price-earnings ratio predicts increasing stock volatility. This result is in line with the findings of Campbell et al. (2018). Because a high price-earnings ratio is empirically found to predict lower returns in the future, this result is again consistent with the present value model. Our finding is also in line with the view that an extreme price-earnings ratio indicates an overvalued market and a subsequent price decline (see Campbell and Shiller, 2001). Similarly, a higher NVIX predicts higher stock volatility in the future. This result squares with the findings in Mittnik et al. (2015) or Conrad and Kleen (2018), among others. The term spread is well known to predict output growth (see, for example, Estrella and Mishkin, 1997, 1998) and, hence, is negatively

[14] Alternatively, a high dividend-price ratio today could forecast lower dividend payments in the future. However, the empirical evidence in Cochrane (1992), among others, suggests that historically the variation in the dividend-price ratio almost exclusively relates to changes in expected excess returns.

Figure 2.7: Number of significant estimates of λ_i for the 49 industry portfolios ($RV_{i,t+1}$ and $\widetilde{RV}_{i,t+1}$, additional predictor variables)



Notes: The plot depicts the number of significantly positive (blue bars) and negative estimates (red bars) of λ_i from the regression $\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \lambda_i w_t + \xi_{i,t+1}$ for the 49 industry portfolios based on predictors $w_t \in \{dpr_t, per_t, ts_t, nvix_t, cay_t, ics_t\}$ when either realized volatility ($RV_{i,t+1}$, left panel) or volatility-adjusted realized volatility ($\widetilde{RV}_{i,t+1}$, right panel) are considered as the volatility proxy. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

related to volatility.

As an intermediate step, we check to what extent the additional predictors are related to expected macroeconomic conditions. We thus regress the forecasts, $x_{t+1|t}$, of the macroeconomic variables on the predictors w_t. The results are reported in Table 2.4.[15]

Table 2.4 shows very reasonable relationships. For example, a higher dividend-price ratio is associated with the expectation of more economic activity and higher rates of inflation. Further, a decrease in the term spread is accompanied by lower growth forecasts and higher inflation forecasts. The former result is in line with the empirical evidence that an inverted yield curve predicts future recessions (see Estrella and Mishkin, 1998). The latter result is consistent with the view that the term spread contains information about the stance of monetary policy (see Estrella and Mishkin, 1997): A central bank responds to increasing inflation expectations by raising short-term interest rates which reduces the term spread.[16] The positive relation between GDP/IP growth expectations and *ics* is in line with the finding in Barsky and Sims (2012) that innovations in consumer confidence contain information about future economic activity. Note that the additional predictor variables explain at maximum 52% of the variation in GDP/IP growth and unemployment

[15] Because the dividend-price ratio and the price-earnings ratio are highly correlated (−0.96), we include them separately in order to avoid multicollinearity.

[16] As discussed in Estrella and Mishkin (1997), the latter effect depends on the credibility of the central bank.

Table 2.4: Regressions of macroeconomic expectations on other predictors

	Dependent variable: Macroeconomic forecasts $x_{t+1\|t}$							
Predictor w_t	Δgdp		Δip		inf		Δune	
dpr	0.47**		0.96**		1.31***		−0.04	
	(0.23)		(0.40)		(0.22)		(0.04)	
per		−0.57**		−1.17***		−1.43***		0.05
		(0.22)		(0.40)		(0.19)		(0.03)
ts	0.94***	0.93***	1.66***	1.64***	−0.70***	−0.75***	−0.07**	−0.07*
	(0.21)	(0.19)	(0.38)	(0.41)	(0.17)	(0.13)	(0.03)	(0.04)
$nvix$	−0.50***	−0.49***	−0.85***	−0.82***	−0.23*	−0.23*	0.07**	0.07*
	(0.10)	(0.10)	(0.22)	(0.19)	(0.12)	(0.13)	(0.04)	(0.04)
cay	−0.64**	−0.64***	−0.92*	−0.94**	−0.01	0.01	0.06**	0.07**
	(0.26)	(0.24)	(0.48)	(0.44)	(0.17)	(0.10)	(0.03)	(0.03)
ics	1.09***	1.16***	1.35**	1.51***	−0.32*	−0.20	−0.19***	−0.21***
	(0.31)	(0.29)	(0.53)	(0.58)	(0.19)	(0.19)	(0.04)	(0.04)
Constant	0.55	−0.38	4.46	2.59	18.77***	16.08***	0.76**	0.83***
	(1.34)	(1.57)	(3.59)	(2.93)	(2.31)	(1.62)	(0.33)	(0.30)
Observations	190	190	190	190	190	190	189	189
$\overline{R}^2$	0.50	0.52	0.30	0.33	0.78	0.82	0.28	0.29

Notes: This table displays the estimates of the regression $x_{t+1|t} = \delta_0 + \delta_1 dpr_t + \delta_2 ts_t + \delta_3 nvix_t + \delta_4 cay_t + \delta_5 ics_t + \zeta_{t+1|t}$ for the forecasts of each macroeconomic variable, i.e., $x_{t+1|t}$. In the even columns, we replace dpr_t with per_t in order to avoid multicollinearity. All predictors are standardized with respect to their standard deviation. The estimation sample $t = 1, \ldots, 190$ covers the period 1968Q4–2016Q1. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. Asterisks *, **, and *** indicate significance at the 10%, 5% and 1% critical level, respectively.

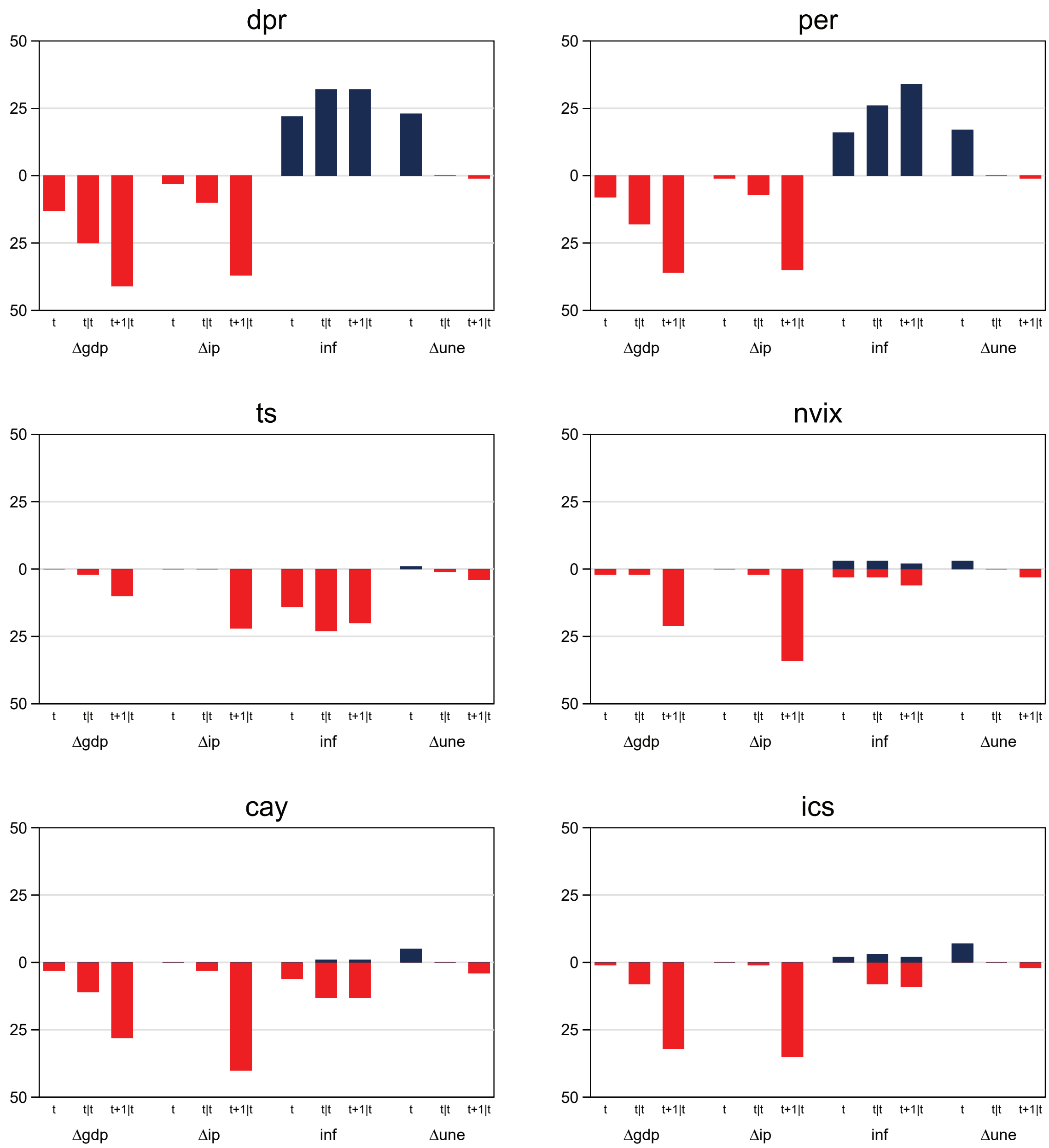
rate forecasts. However, up to 82% of the variation in the inflation forecasts can be explained by the additional predictor variables, whereby the price-earnings ratio and the term spread are the dominant drivers of expected inflation.

Next, we rerun the predictive regressions from Sections 2.6.1 and 2.6.2 and include both X_t as well as the additional predictors (one at a time) as described in Eqn. (2.14). The plots in Figures 2.8 and 2.9 show the outcomes conditional on the six predictor variables (using either $RV_{i,t+1}$ or $\widetilde{RV}_{i,t+1}$ as the proxy variable).

Interestingly, the predictive power of the macroeconomic variables often appears to further increase when including the control variables. For example, when controlling for the dividend-price ratio, the forecasts of GDP/IP growth are significantly related to $RV_{i,t+1}$ in 41/37 cases in Figure 2.8 compared to only 24/33 cases in Figure 2.4 (left panel). A simple explanation could be that controlling for the additional predictors, we obtain more precise estimates of the effects of the macroeconomic variables.

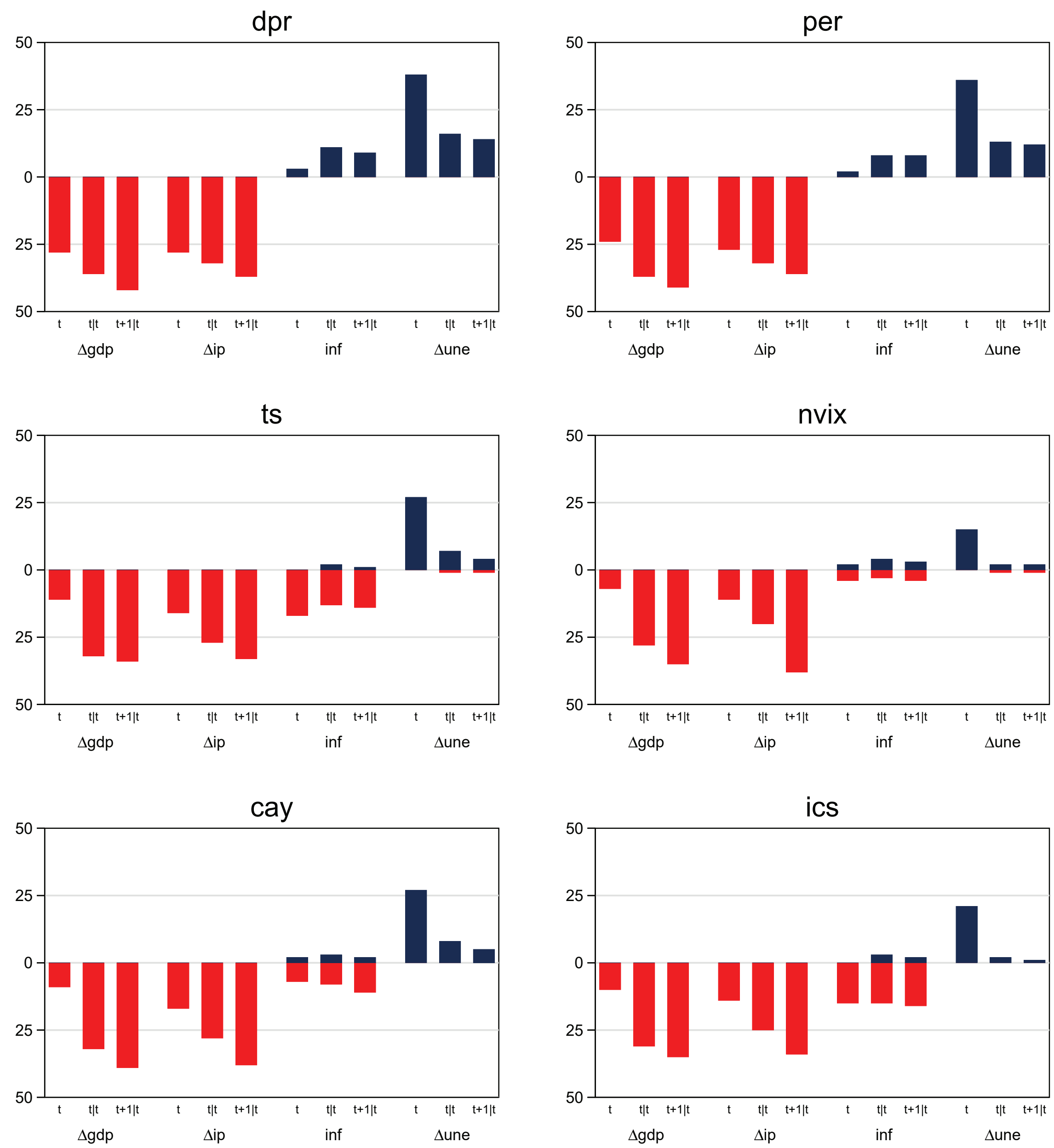
There are only two major changes to our previous findings. First, in many industries the inflation rate has a positive effect for realizations, nowcasts and forecasts when controlling for the dividend-price ratio or the price-earnings ratio. An explanation could be that—when holding the dividend-price/price-earnings ratio constant—an increase in inflation

Figure 2.8: Number of significant estimates of θ_i for the 49 industry portfolios ($RV_{i,t+1}$, controls included)



Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(RV_{i,t+1}) = \phi_{0,i} + \phi_{1,i} \ln(RV_{i,t}) + \phi_{2,i} \ln(RV_{i,t-1}) + \theta_i X_t + \lambda_i w_t + \xi_{i,t+1}$ for the 49 industry portfolios based on predictors $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ and $w_t \in \{dpr_t, per_t, ts_t, nvix_t, cay_t, ics_t\}$. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

Figure 2.9: Number of significant estimates of θ_i for the 49 industry portfolios ($\widetilde{RV}_{i,t+1}$, controls included)



Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(\widetilde{RV}_{i,t+1}) = \phi_{0,i}+\phi_{1,i}\ln(\widetilde{RV}_{i,t})+\phi_{2,i}\ln(\widetilde{RV}_{i,t-1})+\theta_i X_t+\lambda_i w_t+\xi_{i,t+1}$ for the 49 industry portfolios based on predictors $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ and $w_t \in \{dpr_t, per_t, ts_t, nvix_t, cay_t, ics_t\}$. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

is very likely to be associated with a decreasing term spread. Because a decreasing term spread often forecasts a recession, investors may fear a stagflation period which goes along with higher stock volatility. Alternatively, a lower term spread could be the result of (the expectation of) tighter monetary policy and, hence, aligned with higher volatility due to the discount rate effect. To the contrary, when controlling for the term spread, all measures of inflation significantly reduce volatility in around 20 industries. Holding the term spread fixed, inflation is likely to comove with the dividend-price/price-earnings ratio. Higher inflation in combination with a higher dividend-price ratio/a lower price-earnings ratio goes along with decreasing stock volatility due to lower earnings uncertainty. Second, the realized unemployment rate is positively and strongly related to financial volatility when controlling for the dividend-price/price-earnings ratio. Since neither the dividend-price nor the price-earnings ratio are significantly related to the realized unemployment rate (estimates not presented), it appears that controlling for those variables simply reduces the standard error of the estimated effect of the realized unemployment rate.

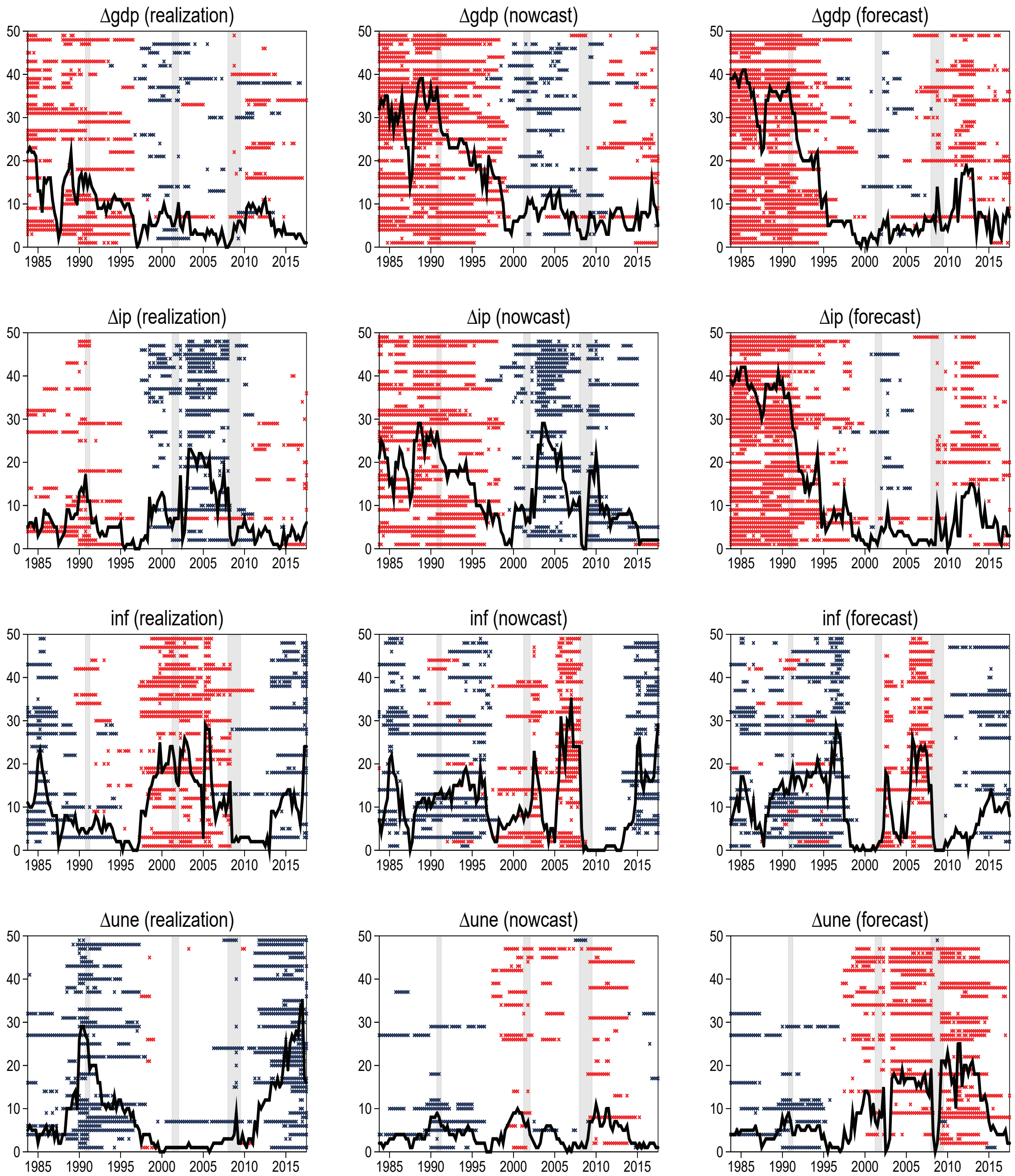
In summary, the macroeconomic variables' predictive ability does not fade away once we control for other standard predictors of returns and volatility. In many cases, the predictive power even increases.

2.6.4 Intertemporal Stability of the Macro-Volatility Nexus

Next, we investigate the stability of the relationship between macroeconomic conditions and financial volatility over time. This is accomplished by estimating a rolling-window version of Eqn. (2.3) for a fixed sample size of 15 years (i.e., 60 quarterly observations). The employed volatility proxy is either industry-specific realized volatility, $RV_{i,t+1}$, or volatility-adjusted industry-specific realized volatility, $\widetilde{RV}_{i,t+1}$. The results over the full sample period are presented in Figures 2.10 and 2.11. We denote by $\theta_{i,t}$ the estimated coefficient on X_t in the 15-year sample that ends at the respective point in time given on the horizontal axis, i.e., quarter t. In each panel, the 49 industry portfolios are depicted on the vertical axis (again sorted by their systematic risk: As can be seen in Table 2.6 in the Appendix, industry '1' (*Precious Metals*) has the lowest and industry '49' (*Business Services*) the highest SR_i). A blue/red cross indicates that for the specific industry the estimate of $\theta_{i,t}$ is significantly positive/negative (at the 5% level). Insignificant estimates are not presented. Finally, the bold black lines indicate for how many industries the estimates of $\theta_{i,t}$ in each window are significant.

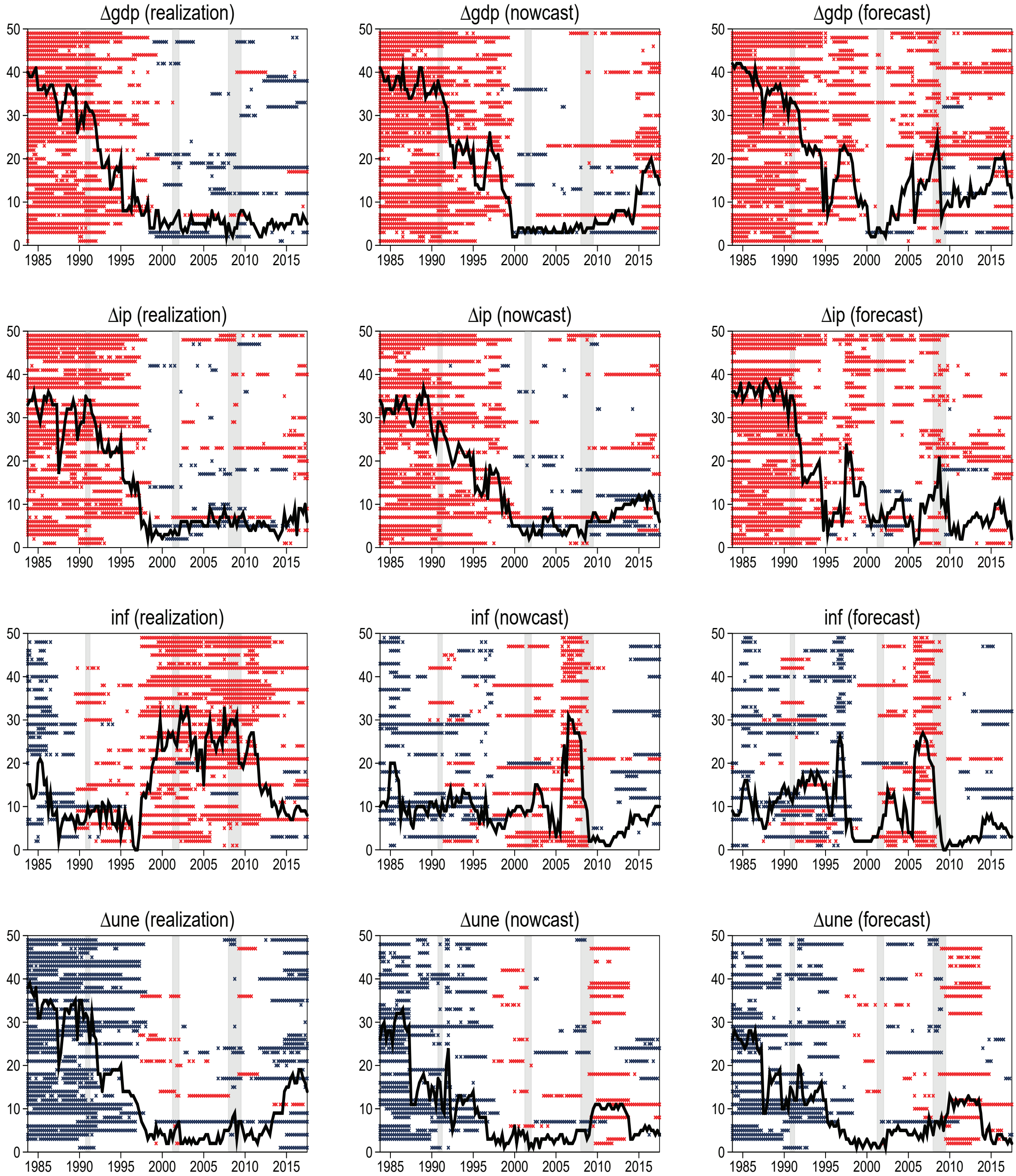
The first row of Figures 2.10/2.11 presents the results for GDP growth realizations (left), nowcasts (middle) and one-step-ahead forecasts (right). The figures clearly show

Figure 2.10: Rolling-window estimates of $\theta_{i,t}$ for the 49 industry portfolios ($RV_{i,t+1}$)



Notes: The plots indicate significantly positive (blue crosses) and negative estimates (red crosses) of $\theta_{i,t}$ from rolling-window regressions $\ln(RV_{i,t+1}) = \phi_{0,i,t} + \phi_{1,i,t}\ln(RV_{i,t}) + \phi_{2,i,t}\ln(RV_{i,t-1}) + \theta_{i,t}X_t + \nu_{i,t+1}$ with window size $\{t-59,\ldots,t\}$ for the 49 industry portfolios, where $RV_{i,t+1}$ denotes realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization (first column), nowcast (second column) or forecast (third column) of the respective macroeconomic variable. The significance level is 5%. Empty spaces indicate insignificant estimates. The bold black lines indicate for how many industries the estimates of $\theta_{i,t}$ in each window are significant. Industries are listed in increasing order in accordance with SR_i. The predictor X_t is standardized with respect to its (full-sample) standard deviation. The estimation sample covers the period 1968Q4–2017Q4 (i.e., the first window ends in 1983Q4). Coefficients are estimated with OLS. We use the variance-covariance estimator by Newey-West (1987), which accounts for arbitrary levels of heteroskedasticity and autocorrelation in the data. Shaded gray areas indicate NBER-based recession periods.

Figure 2.11: Rolling-window estimates of $\theta_{i,t}$ for the 49 industry portfolios ($\widetilde{RV}_{i,t+1}$)



Notes: The plots indicate significantly positive (blue crosses) and negative estimates (red crosses) of $\theta_{i,t}$ from rolling-window regressions $\ln(\widetilde{RV}_{i,t+1}) = \phi_{0,i,t} + \phi_{1,i,t}\ln(\widetilde{RV}_{i,t}) + \phi_{2,i,t}\ln(\widetilde{RV}_{i,t-1}) + \theta_{i,t}X_t + \nu_{i,t+1}$ with window size $\{t-59,\ldots,t\}$ for the 49 industry portfolios, where $\widetilde{RV}_{i,t+1}$ denotes volatility-adjusted realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization (first column), nowcast (second column) or forecast (third column) of the respective macroeconomic variable. The significance level is 5%. Empty spaces indicate insignificant estimates. The bold black lines indicate for how many industries the estimates of $\theta_{i,t}$ in each window are significant. Industries are listed in increasing order in accordance with SR_i. The predictor X_t is standardized with respect to its (full-sample) standard deviation. The estimation sample covers the period 1968Q4–2017Q4 (i.e., the first window ends in 1983Q4). Coefficients are estimated with OLS. We use the variance-covariance estimator by Newey-West (1987), which accounts for arbitrary levels of heteroskedasticity and autocorrelation in the data. Shaded gray areas indicate NBER-based recession periods.

that the nowcasts and one-step-ahead forecasts are more often significant than the realizations. In addition, they reveal that the relevance of GDP growth expectations has changed over time. GDP growth nowcasts/forecasts are statistically significant for almost all industries with a negative estimate of $\theta_{i,t}$ for subsamples ending before 1995. Note that for those early subsamples the average systematic risk was high (see Figure 2.2). For samples that end after 1995, the estimates of $\theta_{i,t}$ become insignificant in most cases and for a few industries even positive for samples that end around 2007. The latter effect is visible in Figure 2.10 but not in Figure 2.11. For more recent subsamples, the estimated coefficients are again significantly negative but in fewer industries than at the beginning of our sample. The observation that GDP growth was a more powerful predictor of volatility during the seventies and eighties was already made in Paye (2012) for the broad stock market. However, the finding that the relevance of GDP growth forecasts has increased recently is new. The recent samples with significant estimates include the Great Recession while the early samples include the oil price shocks of the turbulent 1970s. For both phases, it is reasonable to assume that decreasing GDP growth forecasts/realizations (see Figure 2.3) went along with increasing cash flow uncertainty and, hence, more volatile returns.

The second row of Figure 2.10 presents very similar evidence for IP growth. The main difference is that during the 2000s we find that higher IP growth realizations/nowcasts are often associated with high, rather than low, volatility. Again, this effect is hardly present in Figure 2.11. Since $\widetilde{RV}_{i,t+1}$ is a more accurate measure of the long-term trend in volatility, the positive estimates in Figure 2.10 may be driven by factors that are related to IP growth and that are associated with fluctuations in short-term volatility. We discuss this issue more closely in Section 2.7.2.

As expected, we observe a time-varying relation for inflation. For many industries, nowcasts/forecasts of higher inflation rates tend to increase stock volatility in subsamples that end before the mid-1990s and also in very recent samples. For the turbulent 1970s, our finding is consistent with the stagflation argument of David and Veronesi (2013) and in the 1980s with a hawkish monetary policy during the Volcker disinflation and thereafter. Similarly, during recent years, forecasts of higher inflation rates are likely to increase the probability of a monetary policy tightening and, hence—via the discount effect—predict higher volatility.

In sharp contrast, for samples that end just before the financial crisis of 2007/8, the estimates of $\theta_{i,t}$ for inflation are significantly negative for many industries. That is, before the financial crisis higher inflation rates were associated with lower industry volatility. The empirical results in David and Veronesi (2013) suggest that the 2000s are characterized by deflationary fears and, therefore, higher inflation expectations are good news and accompanied with lower stock volatility.

The results for the unemployment rate are of particular interest. First, we observe a positive effect of the unemployment realizations for subsamples that end before 1995. Those subsamples include data from the late 1970s/early 1980s and the estimated effect is in line with the one that we obtain for GDP/IP growth. Higher rates of unemployment signal economic downturns and go along with high cash flow uncertainty as well as higher stock volatility. Second, for samples that end after 2012, realizations of the unemployment rate have again a positive effect while forecasts have a negative effect for subsamples that end after 2000. The latter effect is in line with the view that the expectation of an increasing unemployment rate either creates the expectation of a looser monetary policy or reduces the probability of a monetary tightening (see Boyd et al., 2005). Both is good news for financial markets and, hence, stock volatility declines.

2.6.5 Out-of-Sample Predictive Ability

While the analysis in Sections 2.6.1 and 2.6.3 was purely in-sample, we can use the rolling-window estimates from the previous section for an out-of-sample evaluation of the predictive ability of the macroeconomic variables. Since we are interested in predicting the actual level of volatility, we exclusively focus on realized volatility for the analysis presented in this subsection. Based on the rolling-window regressions from Figure 2.10, we calculate one-quarter-ahead out-of-sample predictions. We denote the prediction from the AR(2) benchmark model as $\widehat{\ln(RV)}^{(AR)}_{i,t+1|t}$. For each AR(2) model that is augmented by the regressor X_t, the prediction is denoted by $\widehat{\ln(RV)}^{(X)}_{i,t+1|t}$, where $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$. The out-of-sample mean squared prediction errors are calculated as

$$MSE_{i,AR} = \frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\ln(RV)}^{(AR)}_{i,k+1|k} - \ln(RV)_{i,k+1|k}\right)^2 = \frac{1}{K}\sum_{k=1}^{K} e^2_{i,k,AR} \tag{2.15}$$

and

$$MSE_{i,X} = \frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\ln(RV)}^{(X)}_{i,k+1|k} - \ln(RV)_{i,k+1|k}\right)^2 = \frac{1}{K}\sum_{k=1}^{K} e^2_{i,k,X}, \tag{2.16}$$

where K denotes the number of out-of-sample forecasts. To formally test whether the realizations/nowcasts/forecasts lead to significant forecast improvements relative to the AR(2) benchmark, we employ the Giacomini and White (2006) test of equal predictive ability. This test can be implemented by regressing the difference in the squared forecast errors from the AR(2) and the augmented model with the macro variable X_t, i.e.,

$$\Delta SE_{i,k,X} = e^2_{i,k,AR} - e^2_{i,k,X}, \tag{2.17}$$

on a constant and testing for its significance using heteroskedasticity- and autocorrelation-consistent standard errors.[17] Table 2.5 presents a summary of the test decisions. For each variable $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$, we present the number of industries for which the model including X_t significantly improves upon the AR(2) benchmark at the 5% and 10% significance level.

Table 2.5: Giacomini-White tests for the 49 industry portfolios

	Δgdp			Δip			inf			Δune		
	t	$t\|t$	$t+1\|t$	t	$t\|t$	$t+1\|t$	t	$t\|t$	$t+1\|t$	t	$t\|t$	$t+1\|t$
5%	13	4	7	5	8	10	29	16	17	17	11	21
10%	17	10	14	9	12	19	36	18	20	26	19	25

Notes: For each macroeconomic variable, this table displays the rejections from the Giacomini and White (2006) tests for the 49 industry portfolios, i.e., the number of times $MSE_{i,X}$ is significantly lower than $MSE_{i,AR}$ at the 5% (first row) or 10% critical level (second row). The evaluation sample $k = 1, \ldots, 136$ covers the period 1984Q1–2017Q4.

The table shows that even out-of-sample macroeconomic variables are useful predictors of volatility. For example, for the unemployment rate the Giacomini and White (2006) test rejects the null hypothesis for 26/19/25 industries (at the 10% critical level) when using the realizations/nowcasts/forecasts as the predictor. Again, it is important to highlight that the realizations are not available in real-time. Hence, only the forecasting gains from the nowcasts and forecasts can be realized in real-time.

Following Campbell and Thompson (2008), we calculate the out-of-sample R^2 ($R^2_{i,X,OOS}$) for each industry as

$$R^2_{i,X,OOS} = 1 - \frac{MSE_{i,X}}{MSE_{i,AR}}. \tag{2.18}$$

The out-of-sample R^2 provides a gauge of the economic significance of the improvements over the AR(2) model. Figure 2.12 shows scatterplots of $R^2_{i,X,OOS}$ (in percent) for the

[17] We also considered the Giacomini and White (2006) test based on the QLIKE loss function instead of the MSE and obtained qualitatively similar results. Moreover, we considered the Clark and West (2007) test for the equality of the population mean squared errors. This test adjusts for the larger noise associated with $\widehat{\ln(RV)}^{(X)}_{i,t+1|t}$ (compared to $\widehat{\ln(RV)}^{(AR)}_{i,t+1|t}$) due to the estimation of the additional parameter θ_i in the model that includes X_t. The test results were very similar to the ones from the Giacomini and White (2006) test for the equality of the estimated mean squared errors.

SPF predictions versus $R^2_{i,X,OOS}$ (in percent) for the realizations based on the four macro variables. The left panels compare the models based on nowcasts and realizations and the right panels the models based on forecasts and realizations. Points highlighted in red indicate that the MSE for at least one of the models including X_t is significantly lower than for the AR(2) model at the 5% critical level.

First, the figure shows that for almost all industries the mean squared prediction error based on the predictive regression including X_t is lower than the mean squared prediction error of the AR(2) benchmark model, i.e., $R^2_{i,X,OOS} > 0$. Second, for GDP/IP growth the forecasting gains of the nowcasts/forecasts are often higher than or comparable to the goodness-of-fit gains of the realizations. For inflation and the unemployment rate the realizations generate the largest out-of-sample gains. Both findings are broadly in line with our discussion of the Giacomini and White (2006) tests.

Finally, we investigate how the forecasting gains of the predictive regressions including X_t relative to the benchmark AR(2) evolve over time. Following Paye (2012), we calculate the cumulated $\Delta SE_{i,k,X}$, i.e.,

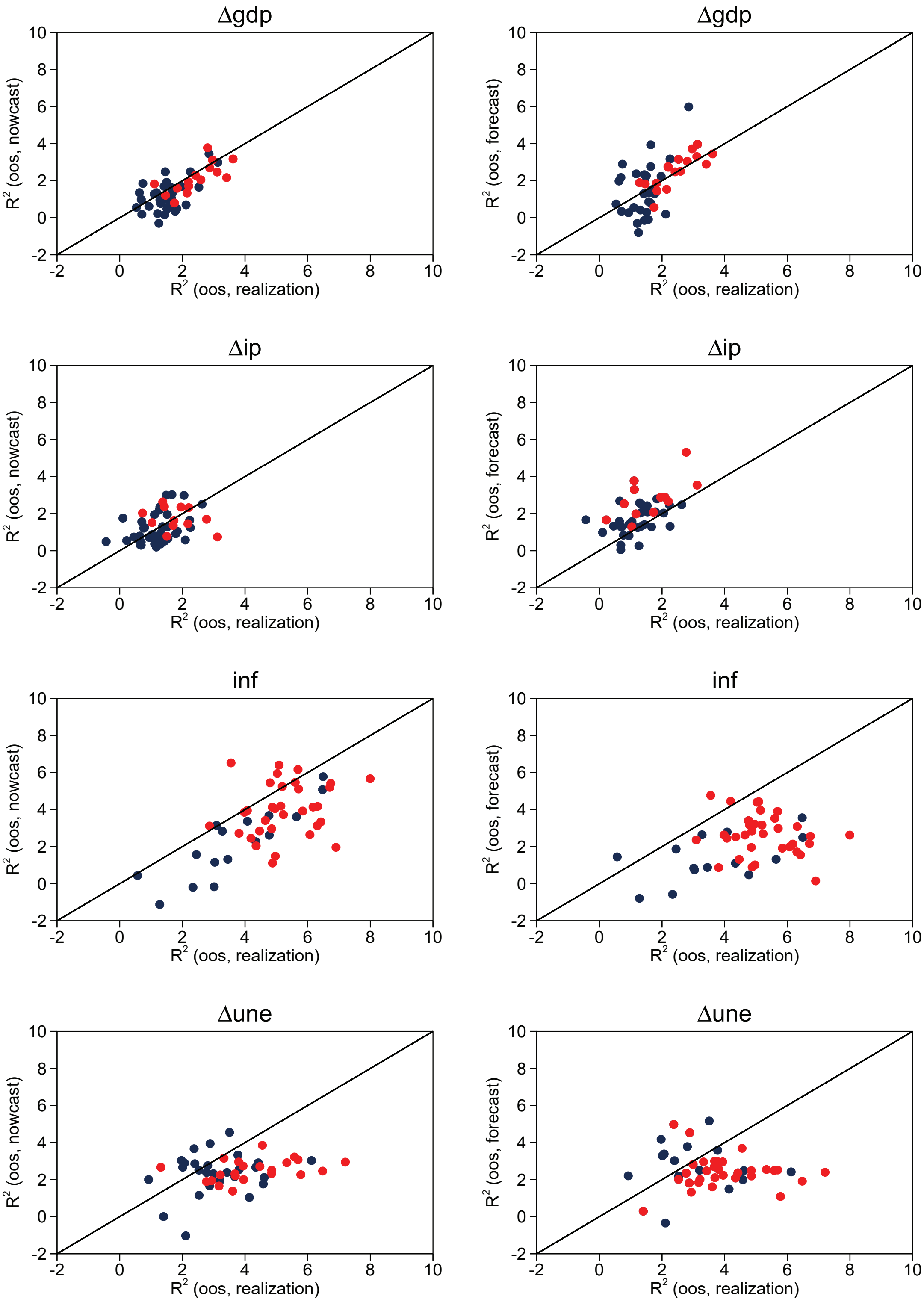
$$cum\Delta SE_{i,\tilde{t},X} = \sum_{k=1}^{\tilde{t}} \Delta SE_{i,k,X}, \tag{2.19}$$

for each industry i and $\tilde{t} = 1, \ldots, K$. For each variable, Figure 2.13 shows the median of the cumulated $\Delta SE_{i,k,X}$ over the 49 industries as well as the 20% and 80% percentiles.

The median cumulated $\Delta SE_{i,k,X}$ is upward trending for all variables, i.e., predictive regressions that are augmented with a macroeconomic variable outperform the benchmark AR(2) essentially over the full out-of-sample period. For expected GDP and IP growth the largest gains are realised during the onset of the Great Recession. Similarly, for all measures of unemployment the largest forecasting gains materialize in the Great Recession and the period thereafter. The biggest gains for inflation are observed for forecasts that are based on rolling-window estimates that do not include the 1970s. That is, when forecasts are based on rolling-window estimates during which the Fed followed a strong inflation objective, the inflation augmented model clearly dominates the benchmark. The fact that the strongest upward trend is observed for realized inflation confirms our previous findings from the Giacomini and White (2006) tests.[18]

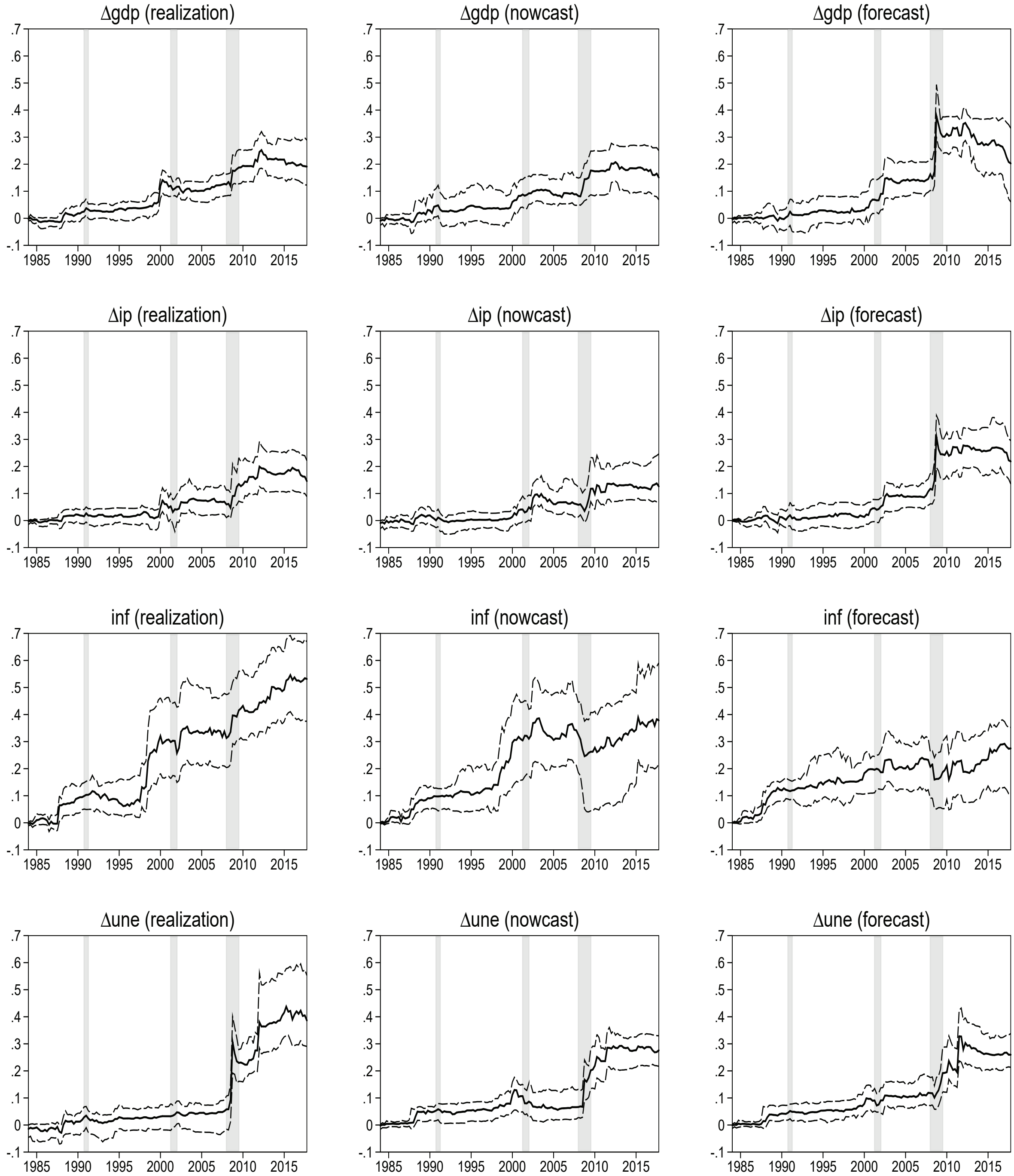
[18] In unreported tests, we have analyzed whether the forecasting gains are evenly spread across the 49 sectors. For inflation and unemployment, the Gini coefficients were close to 0.2. For GDP growth and industrial production we observed a higher inequality of the forecasting gains with Gini coefficients around 0.3.

Figure 2.12: Gain in the out-of-sample goodness of fit for the 49 industry portfolios ($RV_{i,t+1}$)



Notes: The plots in the first column depict the out-of-sample R_i^2 from the 49 industry portfolios when either the nowcast ($x_{t|t}$ vertical axis) or the realization (x_t, horizontal axis) of the respective macroeconomic variable is used as the predictor in the model from Eqn. (2.3) and realized volatility ($RV_{i,t+1}$) is considered as the volatility proxy. In the right column, we replace the out-of-sample R_i^2 for the nowcasts with those for the forecasts. Red dots indicate that the MSE for at least one of the models including X_t is significantly lower than for the AR(2) model at the 5% critical level. The evaluation sample covers the period 1984Q1–2017Q4.

Figure 2.13: Cumulative difference in the squared out-of-sample forecast errors for the 49 industry portfolios ($RV_{i,t+1}$)



Notes: The plots depict the median across the cumulative differences in the squared out-of-sample forecast errors for the 49 industry portfolios, i.e., $\sum_{k=1}^{\tilde{t}}(e^2_{i,k,AR} - e^2_{i,k,X})$ for $i = 1, \ldots, 49$, as a solid black line. The dashed lines represent the corresponding 20% and 80% percentiles. The evaluation sample $k = 1, \ldots, 136$ covers the period 1984Q1–2017Q4.

2.7 Extensions and Robustness

In the following we present several extensions and robustness checks.

2.7.1 Rolling-Window Regressions for the Stock Market

Most of the previous literature has focused on predictive regressions for the broad stock market. To facilitate comparison, we repeat our analysis from Section 2.6.4 for the stock market. Figures 2.14 and 2.15 depict the estimates of $\theta_{m,t}$ when regressing $\ln(RV_{m,t+1})$ or $\ln(\widetilde{RV}_{m,t+1})$ on their own lags and the macroeconomic explanatory variables. As before, each estimate stems from a rolling-window regression with a length of 15 years. Significantly positive/negative estimates are indicated by blue/red dots.

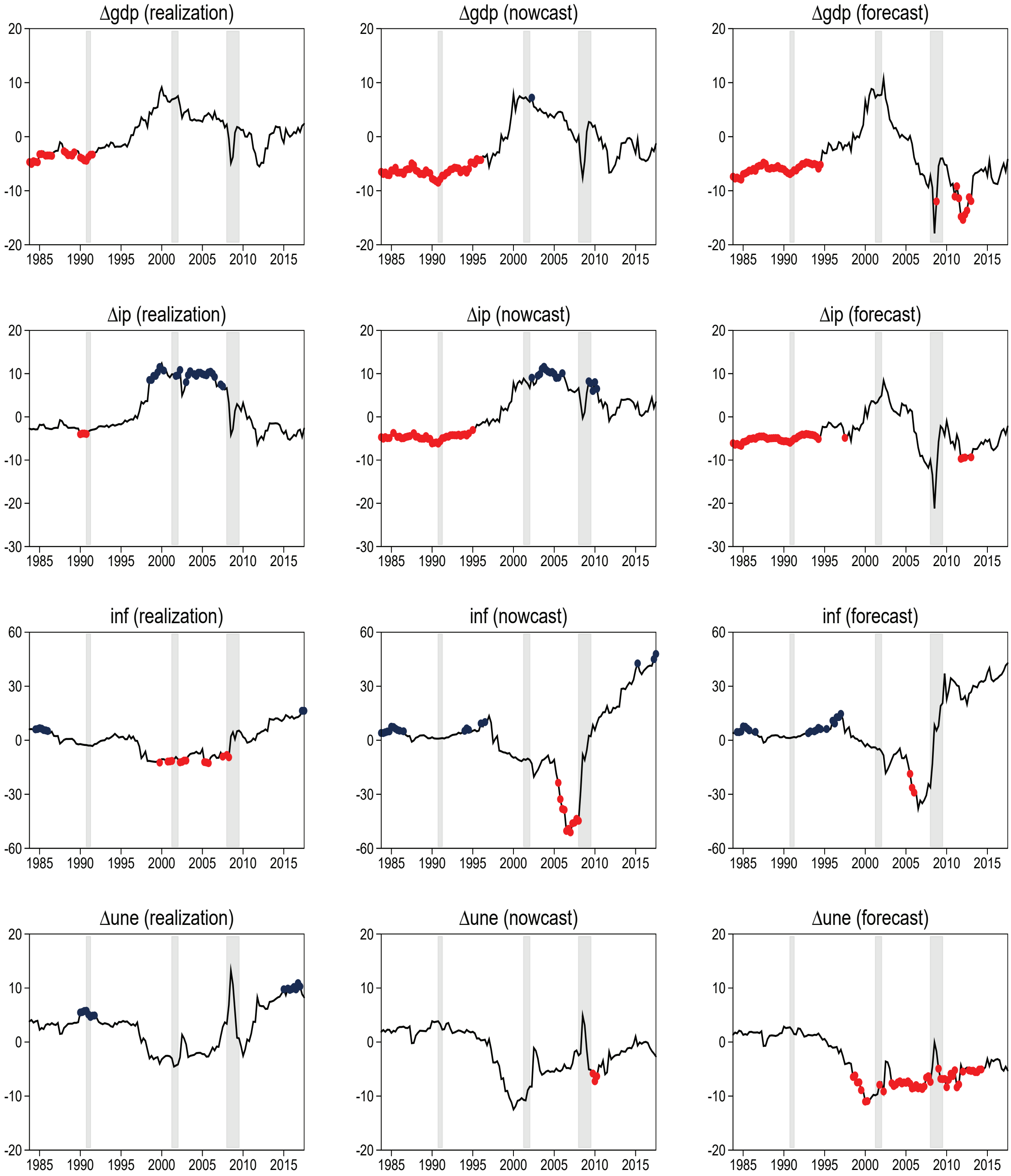
The plots confirm our previous evidence from the 49 industry portfolios. The figures also illustrate why full sample estimates are likely to yield insignificant estimates: The predictive power of the macroeconomic variables is concentrated around specific episodes and, again, we observe switches in the signs of the estimated coefficients.

2.7.2 Rolling-Window Regressions with Price-Earnings Ratio as Control Variable

In this subsection, we investigate more closely why IP growth appears to predict higher rather than lower volatility in subsamples that end between the late 1990s and the Great Recession. This effect is observable in Figures 2.10/2.14 but not in Figures 2.11/2.15 and, hence, may disappear when controlling for variables that correlate with IP and short-term volatility. David and Veronesi (2013, p. 685) argue that in the late 1990s investors believed that the U.S. would enter "into a sustained high-growth regime. These beliefs increased both the price-earnings ratio, as they increased expected cash flows, and volatility, as they increased the uncertainty on whether this transition to a high-growth state was true or not". Hence, we choose the price-earnings ratio as a control variable, i.e., we estimate a rolling-window version of the predictive regression given by Eqn. (2.14) with $RV_{i,t+1}$ as the employed volatility proxy and include the price-earnings ratio as an additional predictor.[19] In this regression, $\theta_{i,t}$ denotes the coefficient on IP growth and $\lambda_{i,t}$ the coefficient on the price-earnings ratio for a window that ends in quarter t. The first row of Figure 2.16 presents the estimates of $\theta_{i,t}$ when controlling for *per* and the second row depicts the estimates of $\lambda_{i,t}$ when controlling for Δip.

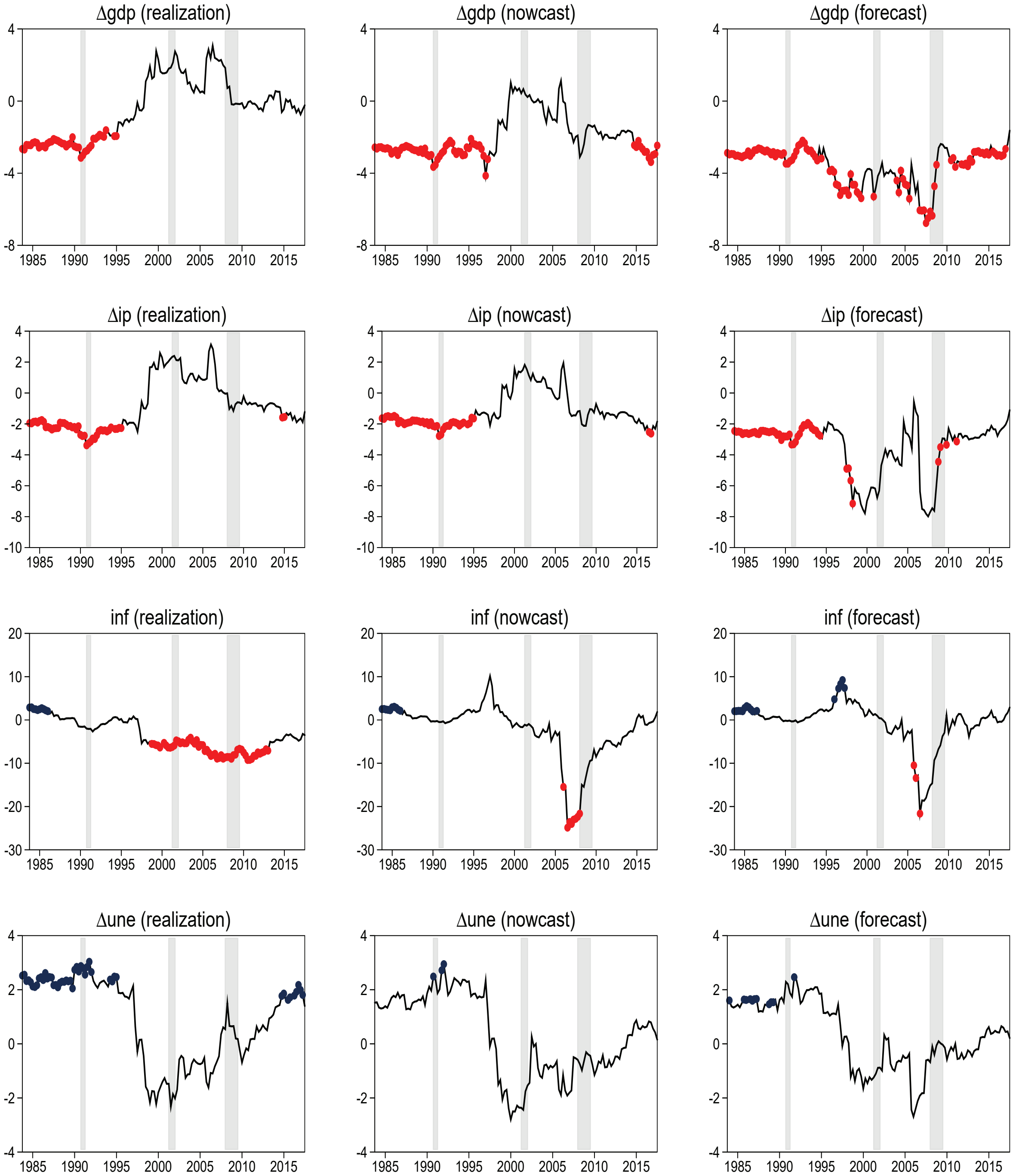
[19] We obtained very similar evidence when controlling for the dividend-price ratio instead of the price-earnings ratio. These results are omitted for reasons of brevity.

Figure 2.14: Rolling-window estimates of $\theta_{m,t}$ $(RV_{m,t+1})$



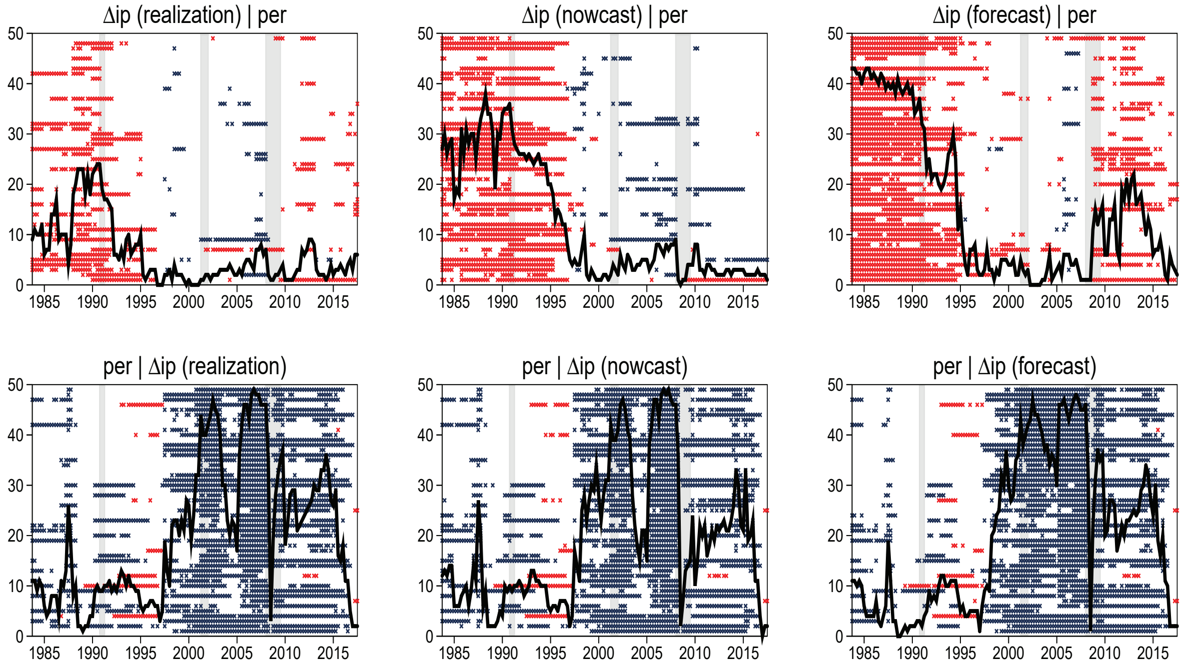
Notes: The plots depict the time series of the estimated slope coefficients $\hat{\theta}_{m,t}$ (times 100) from rolling-window regressions $\ln(RV_{m,t+1}) = \phi_{0,m,t} + \phi_{1,m,t}\ln(RV_{m,t}) + \phi_{2,m,t}\ln(RV_{m,t-1}) + \theta_{m,t}X_t + \nu_{m,t+1}$ with window size $\{t-59,\dots,t\}$ for the market portfolio, where $RV_{m,t+1}$ denotes realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization (first column), nowcast (second column) or forecast (third column) of the respective macroeconomic variable. The constant and the coefficients on the autoregressive terms are not reported. The predictor X_t is standardized with respect to its (full-sample) standard deviation. The estimation sample covers the period 1968Q4–2017Q4 (i.e., the first window ends in 1983Q4). Coefficients are estimated with OLS. We use the variance-covariance estimator by Newey-West (1987), which accounts for arbitrary levels of heteroskedasticity and autocorrelation in the data. Blue dots indicate positive coefficients that are significantly positive at the 5% critical level. Red dots indicate significantly negative estimates. Shaded gray areas indicate NBER-based recession periods.

Figure 2.15: Rolling-window estimates of $\theta_{m,t}$ ($\widetilde{RV}_{m,t+1}$)



Notes: The plots depict the time series of the estimated slope coefficients $\hat{\theta}_{m,t}$ (times 100) from rolling-window regressions $\ln(\widetilde{RV}_{m,t+1}) = \phi_{0,m,t} + \phi_{1,m,t}\ln(\widetilde{RV}_{m,t}) + \phi_{2,m,t}\ln(\widetilde{RV}_{m,t-1}) + \theta_{m,t}X_t + \nu_{m,t+1}$ with window size $\{t-59,\ldots,t\}$ for the market portfolio, where $\widetilde{RV}_{m,t+1}$ denotes volatility-adjusted realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization (first column), nowcast (second column) or forecast (third column) of the respective macroeconomic variable. The constant and the coefficients on the autoregressive terms are not reported. The predictor X_t is standardized with respect to its (full-sample) standard deviation. The estimation sample covers the period 1968Q4–2017Q4 (i.e., the first window ends in 1983Q4). Coefficients are estimated with OLS. We use the variance-covariance estimator by Newey-West (1987), which accounts for arbitrary levels of heteroskedasticity and autocorrelation in the data. Blue dots indicate positive coefficients that are significantly positive at the 5% critical level. Red dots indicate significantly negative estimates. Shaded gray areas indicate NBER-based recession periods.

Figure 2.16: Rolling-window estimates of $\theta_{i,t}$ and $\lambda_{i,t}$ for the 49 industry portfolios when controlling for *per* $(RV_{i,t+1})$



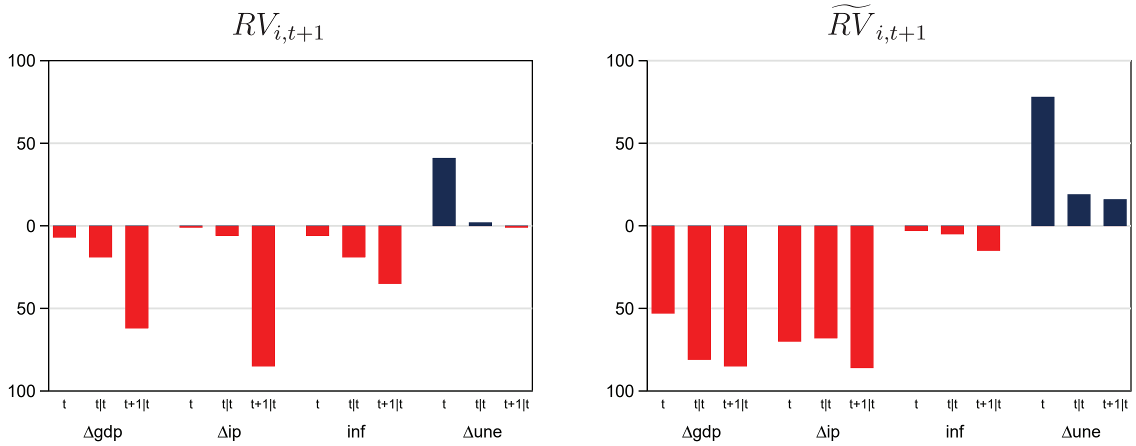
Notes: The plots in the first row indicate the significantly positive (blue crosses) and negative estimates (red crosses) of $\theta_{i,t}$ when controlling for *per* from rolling-window regressions $\ln(RV_{i,t+1}) = \phi_{0,i,t} + \phi_{1,i,t}\ln(RV_{i,t}) + \phi_{2,i,t}\ln(RV_{i,t-1}) + \theta_{i,t}X_t + \lambda_{i,t}per_t + \xi_{i,t+1}$ with window size $\{t-59,\ldots,t\}$ for the 49 industries sample, where $RV_{i,t+1}$ denotes realized volatility, $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization (first column), nowcast (second column) or forecast (third column) of Δip and *per* is the price-earnings-ratio. The plots in the second row depict the corresponding estimates of $\lambda_{i,t}$ when controlling for Δip. The significance level is 5%. Empty spaces indicate insignificant estimates. The bold black lines indicate for how many industries the estimates of $\theta_{i,t}$ and $\lambda_{i,t}$ in each window are significant. Industries are listed in increasing order in accordance with SR_i. The predictor X_t is standardized with respect to its (full-sample) standard deviation. The estimation sample covers the period 1968Q4–2017Q4 (i.e., the first window ends in 1983Q4). Coefficients are estimated with OLS. We use the variance-covariance estimator by Newey-West (1987), which accounts for arbitrary levels of heteroskedasticity and autocorrelation in the data. Shaded gray areas indicate NBER-based recession periods.

As the upper row of the figure shows, almost all of the significantly positive estimates of $\theta_{i,t}$ disappear when controlling for the price-earnings ratio. This suggests that—without controlling for the price-earnings ratio—high IP growth may proxy a high price-earnings ratio, i.e., lofty market valuations and high uncertainty, and, hence, predict higher volatility. In particular, this appears to be the case for subsamples that include the dot-com bubble. This intuition is confirmed by the lower row of Figure 2.16 which shows the estimates of $\lambda_{i,t}$ which are almost exclusively positive.

2.7.3 Portfolios Formed on Size and Book-to-Market

We also investigate whether our results extend to other types of portfolios. Instead of the 49 industry portfolios, we consider the volatility of 100 portfolios formed on size and book-to-market. The return data are again obtained from the Fama-French Data Library. Figure 2.17 shows the results for full sample predictive regressions with $RV_{i,t}$ (left panel) and $\widetilde{RV}_{i,t}$ (right panel) as the employed proxy for volatility. As for the 49 industry portfolios, in the full sample forecasts of GDP/IP growth and inflation are much more often significant than the corresponding realizations. The opposite is true for the unemployment rate. Again, we find that in predictive regressions based on $\widetilde{RV}_{i,t}$ the realizations and nowcasts are more frequently significant than in regressions based on $RV_{i,t}$. This underlines again that $RV_{i,t}$ is a noisy measure of long-term volatility.

Figure 2.17: Number of significant estimates of θ_i for the 100 portfolios formed on size and book-to-market ($RV_{i,t+1}$ and $\widetilde{RV}_{i,t+1}$)

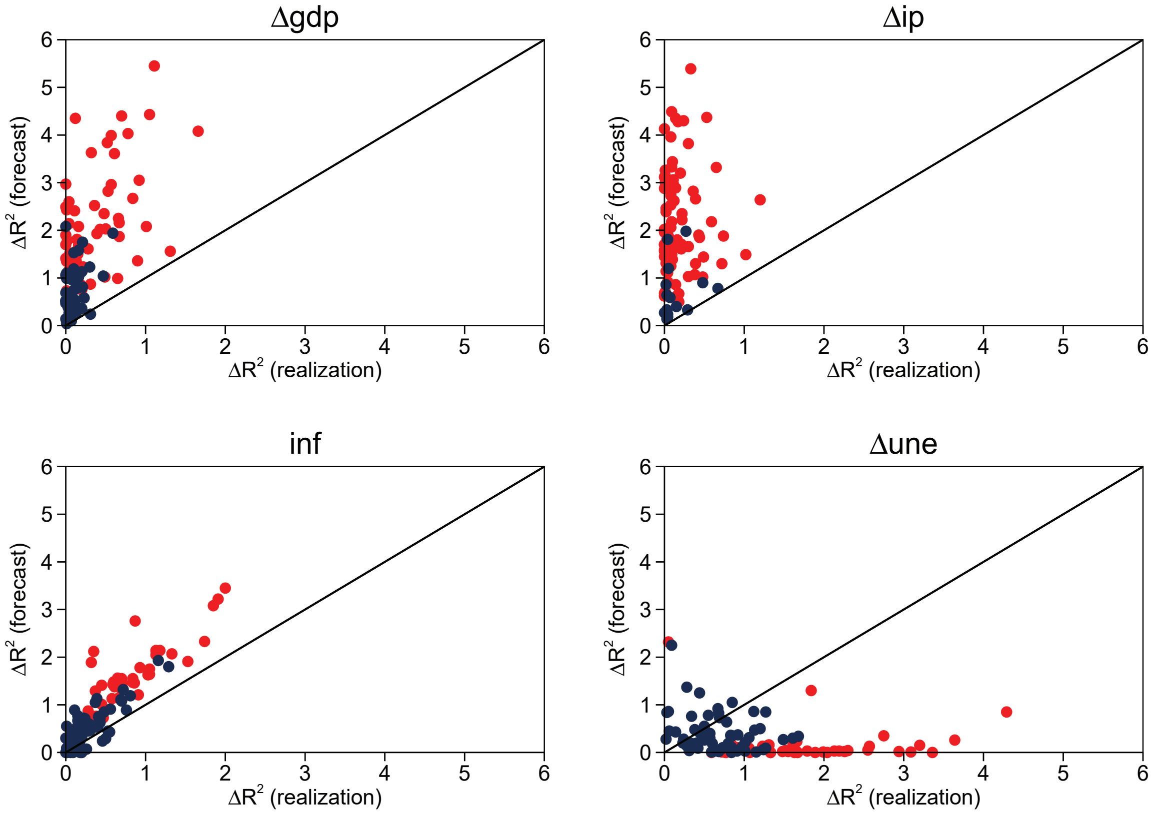


Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the 100 portfolios formed on size and book-to-market based on predictors $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ when either realized volatility ($RV_{i,t+1}$, left panel) or volatility-adjusted realized volatility ($\widetilde{RV}_{i,t+1}$, right panel) are considered as the volatility proxy. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

Figure 2.18 shows the percentage increases in R_i^2 as compared to the pure AR(2). As in Figure 2.5, we compare forecasts and realizations and use $RV_{i,t}$ as the volatility proxy. The results strongly support our previous findings. For GDP/IP growth and inflation the strongest percentage increases are observed for the forecasts rather than the realizations. Again, for the unemployment rate the realizations are more important.

Thus, the results for the industry portfolios directly extend to portfolios formed on size and book-to-market. We also considered 100 portfolios formed on size and either operating

Figure 2.18: Gain in the goodness of fit for the 100 portfolios formed on size and book-to-market ($RV_{i,t+1}$)



Notes: The plots depict the percentage increase in R_i^2 relative to the AR(2) benchmark for the 100 portfolios formed on size and book-to-market when either the forecast ($x_{t+1|t}$ vertical axis) or the realization (x_t, horizontal axis) of the respective macroeconomic variable is considered as the predictor in the model from Eqn. (2.3) and realized volatility ($RV_{i,t+1}$) is considered as the volatility proxy. Red dots indicate that in the underlying regression, θ_i is significantly different from zero at the 5% level either for x_t, $x_{t+1|t}$, or both. The estimation sample covers the period 1968Q4–2017Q4.

profitability or investment and, again, obtained very similar results (not reported).

2.7.4 Idiosyncratic Volatility

Next, we investigate whether our findings hold for idiosyncratic volatility as well. It might well be that the explanatory power of the macroeconomic variables only works through the market factor. In this case, idiosyncratic volatility should be unrelated to macroeconomic conditions. Based on the residuals from Eqn. (2.6), $\hat{\eta}_{i,d,t} = r_{i,d,t} - \hat{\mu}_i - \hat{\beta}_i r_{m,d,t}$, quarterly idiosyncratic volatility in industry i is calculated as

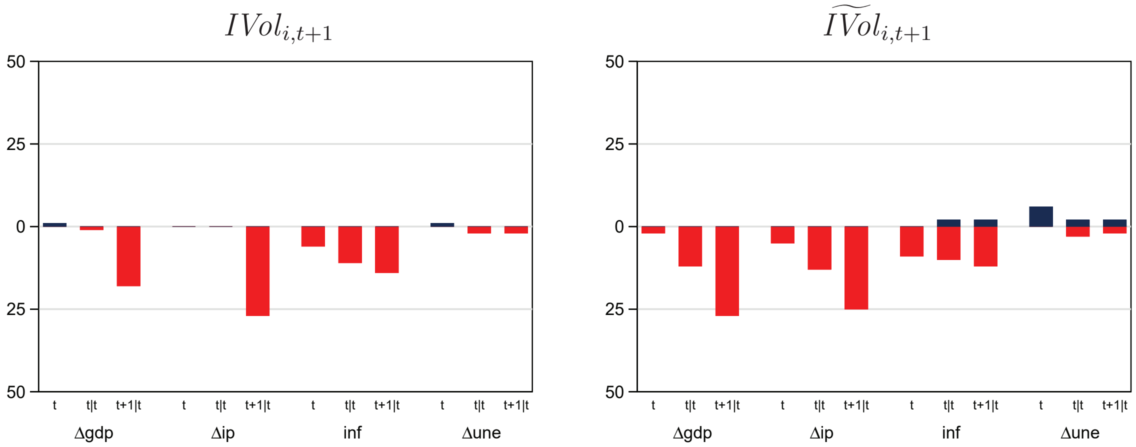
$$IVol_{i,t} = \sqrt{\sum_{d=1}^{D_t} \hat{\eta}_{i,d,t}^2}. \tag{2.20}$$

Following the approach in Section 2.6.2, we also construct a measure of volatility-adjusted idiosyncratic volatility. Using Eqn. (2.6) as the mean equation, we obtain

$$\widetilde{IVol}_{i,t} = \sqrt{\frac{\hat{\omega}_i}{1 - \hat{\alpha}_i - \hat{\delta}_i - \hat{\gamma}_i/2}} \cdot \sqrt{\sum_{d=1}^{D_t} \left(\hat{\eta}_{i,d,t}^2 / \hat{h}_{i,d,t}\right)}. \tag{2.21}$$

Summary statistics for both measures of idiosyncratic volatility for the market and the 5 industry portfolios are reported in Table 2.1. Below, we consider the 49 industry portfolios. Our findings for $IVol_{i,t}$ and $\widetilde{IVol}_{i,t}$ are presented in Figure 2.19. The improvements in the goodness of fit as compared to predictive regressions based on the AR(2) benchmark in the case of $IVol_{i,t}$ are depicted in Figure 2.20.

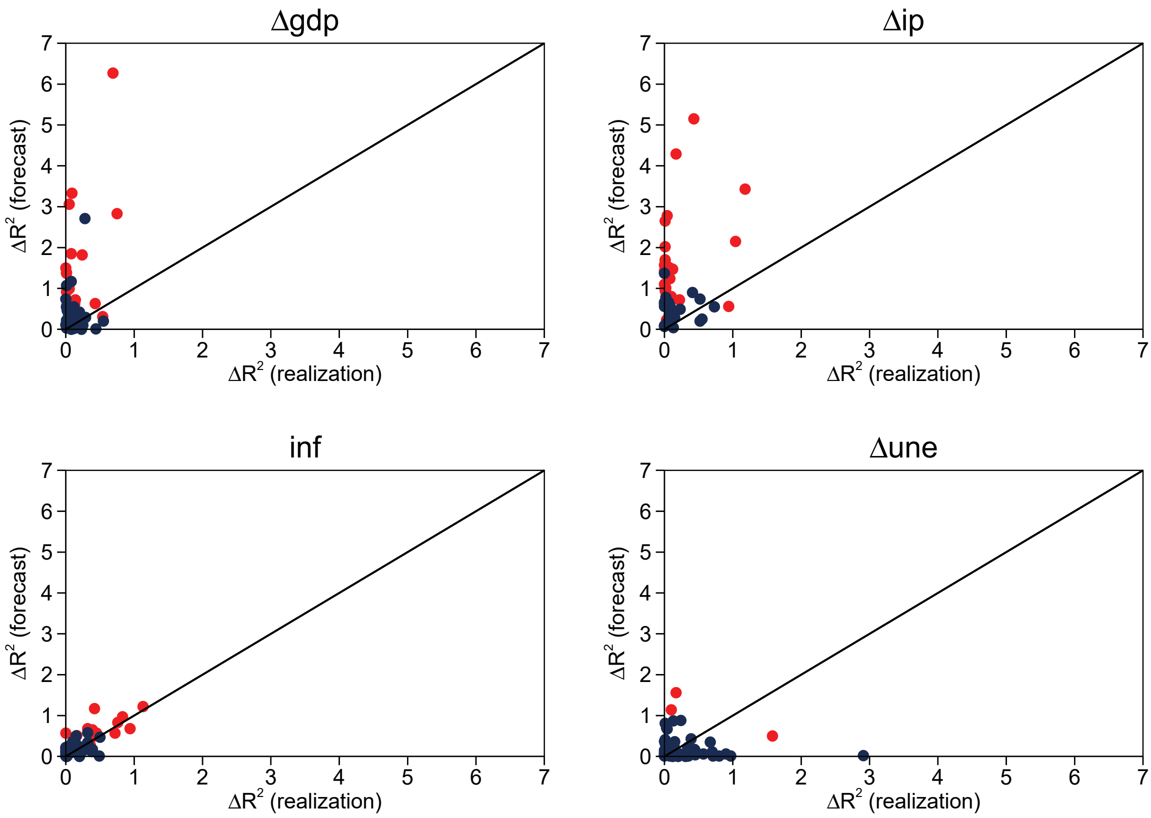
Figure 2.19: Number of significant estimates of θ_i for the 49 industry portfolios ($IVol_{i,t+1}$ and $\widetilde{IVol}_{i,t+1}$)



Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the 49 industry portfolios based on predictors $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ when either idiosyncratic volatility ($IVol_{i,t+1}$, left panel) or volatility-adjusted idiosyncratic volatility ($\widetilde{IVol}_{i,t+1}$, right panel) are considered as the volatility proxy. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

Interestingly, the number of significant estimates of θ_i is only marginally smaller than in Figure 2.5 for realized volatility. Thus, the market factor alone does not capture the relation between macroeconomic conditions and industry-specific volatility. Engle and Rangel (2012) question the standard CAPM assumption that $\mathbf{Cov}(\eta_{i,d,t}, \eta_{j,d,t}) = 0$

Figure 2.20: Gain in the goodness of fit for the 49 industry portfolios ($IVol_{i,t+1}$)



Notes: The plots depict the percentage increase in R_i^2 relative to the AR(2) benchmark for the 49 industry portfolios when either the forecast ($x_{t+1|t}$ vertical axis) or the realization (x_t, horizontal axis) of the respective macroeconomic variable is considered as the predictor in the model from Eqn. (2.3) and idiosyncratic volatility ($IVol_{i,t+1}$) is considered as the volatility proxy. Red dots indicate that in the underlying regression, θ_i is significantly different from zero at the 5% level either for x_t, $x_{t+1|t}$, or both. The estimation sample covers the period 1968Q4–2017Q4.

for $i \neq j$ and propose the Factor-Spline-GARCH model which allows for comovements in $\eta_{i,d,t}$ and specifies the conditional variance of $\eta_{i,d,t}$ as a two component process. Our findings suggest that $\mathbf{Cov}(\eta_{i,d,t}^2, \eta_{j,d,t}^2) \neq 0$ for $i \neq j$ and that the idiosyncratic long-term volatilities in different industries are driven by the same macroeconomic variables.

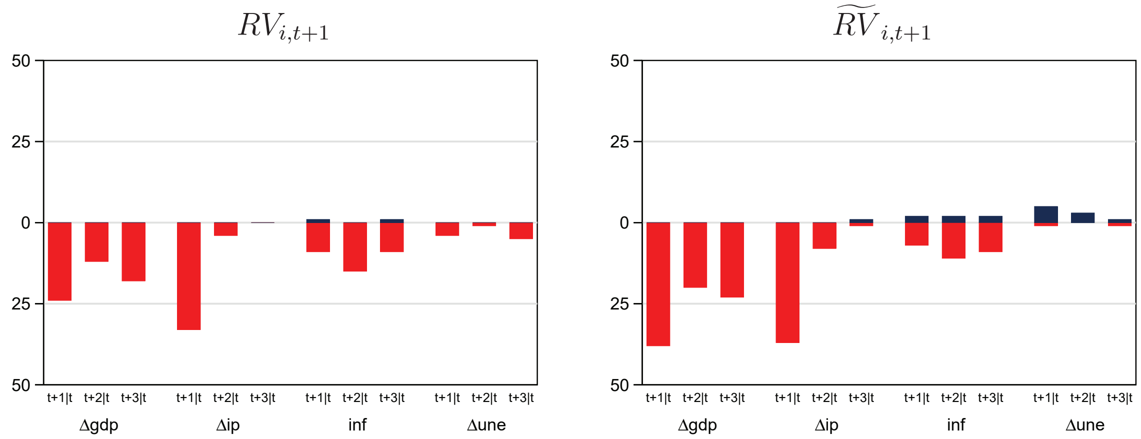
2.7.5 Multi-Step-Ahead Forecasts

Besides the one-quarter-ahead forecasts, $x_{t+1|t}$, the SPF also contains forecasts for two- and three-quarters-ahead (i.e., $x_{t+2|t}$ and $x_{t+3|t}$).[20] Since we exclusively focused on $x_{t+1|t}$ so far, checking for the predictive power of $x_{t+2|t}$ and $x_{t+3|t}$ appears appropriate. Figure 2.21

[20] The SPF also provides four-quarters-ahead forecasts, but—in particular at the beginning of the sample period—there are quite a few missing observations. Therefore, we refrain from using those.

shows the number of significant θ_i-estimates in the 49 industry portfolios for all three forecasts when either using $RV_{i,t+1}$ (left panel) or $\widetilde{RV}_{i,t+1}$ (right panel) as the proxy variable.

Figure 2.21: Number of significant estimates of θ_i for the 49 industry portfolios ($RV_{i,t+1}$ and $\widetilde{RV}_{i,t+1}$, forecast horizons)



Notes: The plots depict the number of significantly positive (blue bars) and negative estimates (red bars) of θ_i from the regression $\ln(Vol_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(Vol_{i,t}) + \phi_{2,i}\ln(Vol_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the 49 industry portfolios based on predictors $X_t \in \{x_{t+1|t}, x_{t+2|t}, x_{t+3|t}\}$ when either realized volatility ($RV_{i,t+1}$, left panel) or volatility-adjusted realized volatility ($\widetilde{RV}_{i,t+1}$, right panel) are considered as the volatility proxy. The significance level is 5%. The estimation sample covers the period 1968Q4–2017Q4.

The figure clearly shows that for GDP/IP growth the one-quarter-ahead forecasts are most useful. For inflation and unemployment the evidence is more mixed. Nevertheless, Figure 2.21 again confirms the view that forecasts of macroeconomic variables are powerful predictors of future stock volatility.

2.8 Conclusion

We reconsider the question to what extent macroeconomic variables have predictive power for future stock volatility in predictive regressions when controlling for past volatility. The econometric framework of this chapter is closest to Paye (2012, p. 545) who concludes that "only modest forecasting gains are possible" because "volatility comoves tightly with the business cycle and lagged volatility itself contains a wealth of information about business conditions". In contrast, our results suggest that survey forecasts of macroeconomic variables contain information about future volatility that is complementary to the information included in lagged volatility. By considering 49 industry portfolios, we show that higher forecasts of GDP/IP growth decrease industry-level volatility in many sectors in subsamples that include the 1970s and early 1980s. For subsamples that include the period before the Great Recession or the most recent years, realizations of inflation and realizations/forecasts of the unemployment rate are powerful predictors. Our findings highlight that the relation between macroeconomic conditions and stock volatility changes over time and also varies across industries. This may explain the inconclusive results from the previous literature which usually focused on the aggregate stock market.

2.9 Appendix

Table 2.6: Description of the 49 industry portfolios

i	Sector	Description	SR_i	i	Sector	Description	SR_i
49	*bussv*	Business Services	0.81	24	*rubbr*	Rubber and Plastic Products	0.57
48	*mach*	Machinery	0.77	23	*boxes*	Shipping Containers	0.56
47	*fin*	Trading	0.75	22	*meals*	Restaurants, Hotels, Motels	0.56
46	*whlsl*	Wholesale	0.71	21	*fun*	Entertainment	0.55
45	*elceq*	Electrical Equipment	0.71	20	*food*	Food Products	0.54
44	*rtail*	Retail	0.70	19	*persv*	Personal Services	0.52
43	*chems*	Chemicals	0.70	18	*oil*	Petroleum and Natural Gas	0.51
42	*insur*	Insurance	0.69	17	*toys*	Recreation	0.50
41	*bldmt*	Construction Materials	0.69	16	*txtls*	Textiles	0.49
40	*trans*	Transportation	0.69	15	*other*	Almost Nothing	0.49
39	*labeq*	Measuring and Control Equipment	0.68	14	*util*	Utilities	0.48
38	*chips*	Electronic Equipment	0.68	13	*rlest*	Real Estate	0.46
37	*banks*	Banking	0.67	12	*mines*	Non-Metallic and Industrial Metal Mining	0.46
36	*telcm*	Communication	0.66	11	*fabpr*	Fabricated Products	0.43
35	*paper*	Business Supplies	0.65	10	*hlth*	Healthcare	0.39
34	*books*	Printing and Publishing	0.63	9	*beer*	Beer & Liquor	0.37
33	*autos*	Automobiles and Trucks	0.63	8	*softw*	Computer Software	0.36
32	*steel*	Steel Works Etc	0.62	7	*ships*	Shipbuilding, Railroad Equipment	0.35
31	*hardw*	Computers	0.62	6	*soda*	Candy & Soda	0.31
30	*drugs*	Pharmaceutical Products	0.61	5	*agric*	Agriculture	0.30
29	*cnstr*	Construction	0.60	4	*guns*	Defense	0.29
28	*clths*	Apparel	0.60	3	*coal*	Coal	0.29
27	*aero*	Aircraft	0.60	2	*smoke*	Tobacco Products	0.26
26	*hshld*	Consumer Goods	0.58	1	*gold*	Precious Metals	0.03
25	*medeq*	Medical Equipment	0.58				

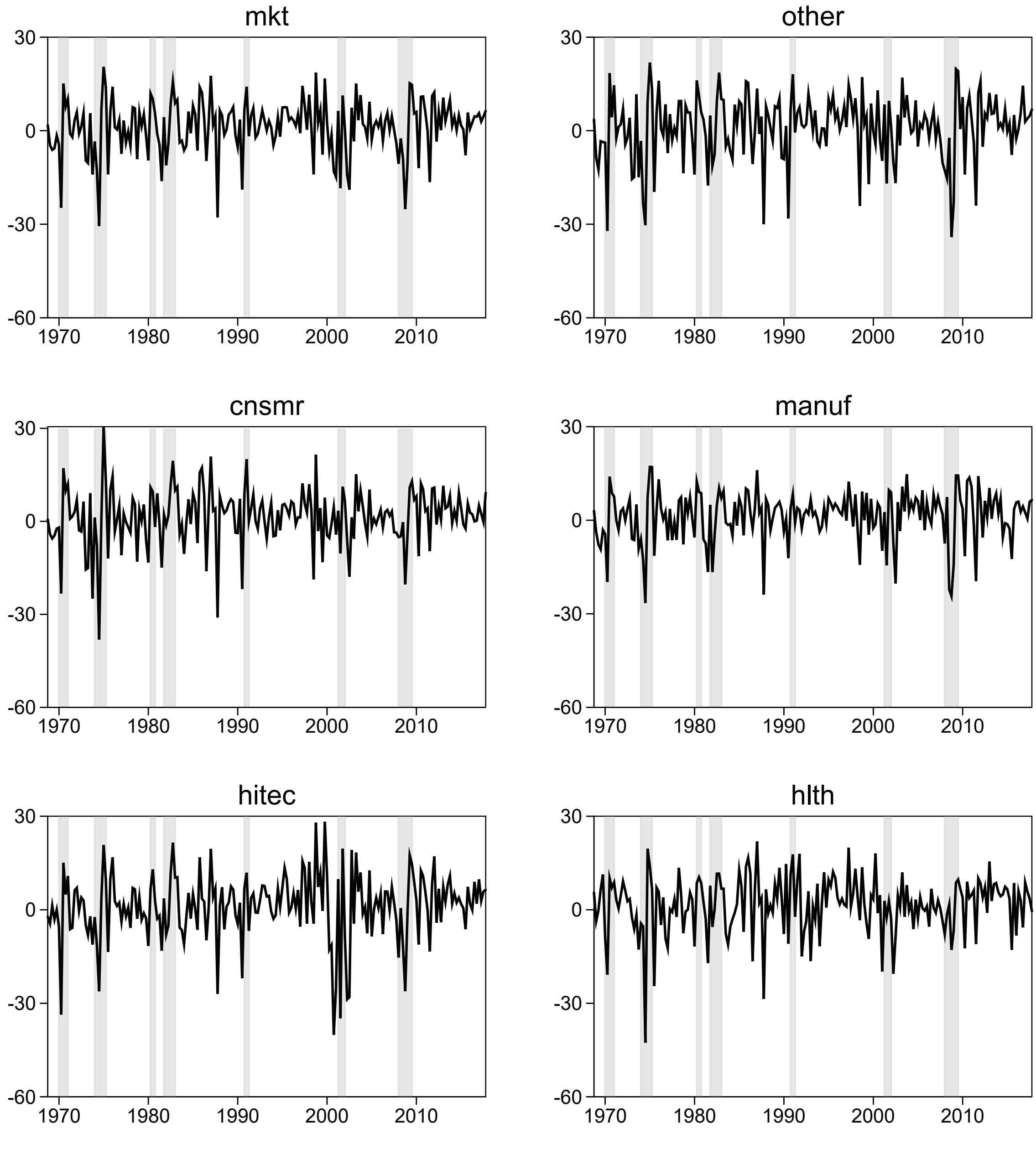
Notes: This table provides a description of the 49 industry portfolios following the definitions from the Fama-French data library. We also report the systematic risk of each portfolio based on the estimates of Eqn. (2.6).

Table 2.7: Predictive regressions for volatility-adjusted realized volatility (market and 5 industry portfolios)

		Industry portfolio											
		mkt		*other*		*cnsmr*		*manuf*		*hitec*		*hlth*	
Predictor x	Timing	$\hat{\theta}_m$	ΔR^2_m	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i	$\hat{\theta}_i$	ΔR^2_i
Δgdp	t	−1.37**	6.94	−1.17*	6.47	−1.03**	4.15	−1.24**	8.36	−1.10	9.62	−0.36	0.84
		(0.60)		(0.65)		(0.51)		(0.54)		(0.68)		(0.65)	
Δgdp	$t\|t$	−2.15***	15.85	−2.15***	20.73	−1.46***	8.09	−1.78***	16.20	−1.53***	17.56	−1.15*	8.03
		(0.42)		(0.68)		(0.34)		(0.53)		(0.52)		(0.60)	
Δgdp	$t+1\|t$	−2.82***	27.98	−2.74***	34.03	−1.90***	14.07	−2.37***	28.43	−2.19***	36.99	−1.53***	14.31
		(0.47)		(0.77)		(0.42)		(0.48)		(0.64)		(0.53)	
Δip	t	−1.31**	6.28	−1.41*	9.18	−0.84	2.74	−1.30**	9.19	−0.95	6.95	−0.39	0.95
		(0.65)		(0.76)		(0.64)		(0.60)		(0.58)		(0.74)	
Δip	$t\|t$	−1.39***	6.85	−1.46**	9.81	−0.61*	1.45	−1.54***	12.64	−0.97**	7.23	−0.65	2.66
		(0.38)		(0.67)		(0.37)		(0.50)		(0.41)		(0.58)	
Δip	$t+1\|t$	−2.68***	26.71	−2.69***	34.29	−1.90***	14.41	−2.22***	26.35	−2.24***	39.82	−1.82***	20.69
		(0.51)		(0.55)		(0.46)		(0.59)		(0.72)		(0.54)	
inf	t	−0.30	0.36	−1.40	9.40	0.14	0.07	−0.10	0.06	−0.89	6.40	−0.35	0.80
		(0.80)		(0.94)		(0.64)		(0.71)		(0.90)		(0.63)	
inf	$t\|t$	0.20	0.15	−0.93	4.01	0.43	0.76	0.30	0.50	−0.28	0.61	0.10	0.07
		(0.61)		(0.61)		(0.60)		(0.63)		(0.73)		(0.56)	
inf	$t+1\|t$	0.23	0.21	−0.98*	4.45	0.41	0.67	0.33	0.63	−0.20	0.33	0.21	0.29
		(0.61)		(0.55)		(0.62)		(0.60)		(0.75)		(0.62)	
Δune	t	1.43**	7.60	1.55***	11.25	1.36**	7.33	1.10**	6.69	1.16*	9.94	0.47	1.57
		(0.55)		(0.55)		(0.62)		(0.55)		(0.68)		(0.71)	
Δune	$t\|t$	0.62	1.41	0.88	3.63	0.36	0.52	0.58	1.87	0.00	0.00	−0.17	0.21
		(0.65)		(0.65)		(0.59)		(0.67)		(0.68)		(0.69)	
Δune	$t+1\|t$	0.66	1.63	0.91	3.94	0.24	0.23	0.59	1.95	0.02	0.00	−0.14	0.14
		(0.54)		(0.64)		(0.55)		(0.67)		(0.76)		(0.73)	

Notes: This table displays the estimated slope coefficients $\hat{\theta}_i$ from the regression $\ln(\widetilde{RV}_{i,t+1}) = \phi_{0,i} + \phi_{1,i}\ln(\widetilde{RV}_{i,t}) + \phi_{2,i}\ln(\widetilde{RV}_{i,t-1}) + \theta_i X_t + \nu_{i,t+1}$ for the market portfolio and the 5 industry portfolios, where $\widetilde{RV}_{i,t+1}$ denotes volatility-adjusted realized volatility and $X_t \in \{x_t, x_{t|t}, x_{t+1|t}\}$ is either the realization, the nowcast or the forecast of the respective macroeconomic variable. The percentage increase in the R^2 relative to the AR(2) benchmark is denoted by ΔR^2. The R^2_m for the market portfolio based on the AR(2) benchmark is 0.18. The constant and the coefficients on the autoregressive terms are not reported. The predictor X_t is standardized with respect to its standard deviation. Industries are listed in decreasing order according to the SR_i from Eqn. (2.6) (see Table 2.1). The estimation sample $t = 1, \ldots, 197$ covers the period 1968Q4–2017Q4. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. The coefficients and standard errors are the estimated ones times 100. Asterisks *, **, and *** indicate significance at the 10%, 5% and 1% critical level, respectively.

Figure 2.22: Quarterly returns for the market and the 5 industry portfolios



Notes: The plots depict the time series of the quarterly value-weighted excess returns, i.e., $r_{i,t}$, for the market (*mkt*) as well as the other (*other*), consumer durables (*cnsmr*), manufacturing (*manuf*), business equipment (*hitec*) and healthcare (*hlth*) industries. Sectors are listed in decreasing order according to the systematic risk from Eqn. (2.6) (see Table 2.1). The data are taken from the Fama-French data library. The sample period is 1968Q4–2017Q4. Shaded gray areas indicate NBER-based recession periods.

Chapter 3

Inflation Uncertainty, Disagreement and Monetary Policy

3.1 Introduction

Inflation uncertainty, as measured by the average variance from a cross-section of individual density forecasts from the European Central Bank's Survey of Professional Forecasters, has considerably increased since at least the beginning of the financial crisis in 2008. The emergence of inflation uncertainty may reflect concerns about further rising inflation due to expansive monetary policy or fears of deflation as a result of a prolonged recession period (Drakos and Kouretas, 2015). Several theoretical arguments suggest that macroeconomic uncertainty has negative welfare effects (Fernández-Villaverde et al., 2011; Bloom et al., 2014). For example, inflation uncertainty may induce agents to postpone investment or saving decisions and reduce market efficiency due to an increase in the volatility of both relative prices and risks regarding income streams from nominal financial and wage contracts (Friedman, 1977; Baillie et al., 1996; Bloom, 2009). Empirical evidence for these arguments is documented, e.g., by Grier and Perry (2000), Grier et al. (2004) and Wright (2011). In contrast, the level of inflation and inflation expectations have been relatively low throughout the last decades and evolve in a rather stable way (Galí and Gambetti, 2009; Lahiri and Sheng, 2010). This is often ascribed to successful monetary policy. However, the increase in inflation uncertainty questions these potential achievements.

This chapter is based on a paper that I wrote jointly with Matthias Hartmann. A similar version of this analysis has been published under the name "Inflation uncertainty, disagreement and monetary policy: Evidence from the ECB Survey of Professional Forecasters" in a special issue of the *Journal of Empirical Finance* on "The Euro Zone in Crisis" (cf. Glas and Hartmann, 2016).

Though survey-based measures of inflation uncertainty (henceforth: IU) such as the average variance from a cross-section of density forecasts (average individual IU) are often regarded as one of the most reliable ways to quantify uncertainty (Bachmann et al., 2013; Clements, 2014), many surveys do not elicit density forecasts. In such cases, the average variance of the point forecasts (disagreement) is often used as a proxy variable for IU. In addition, disagreement itself is often considered as a variable of interest, e.g., due to its potential influence on aggregate output (Mankiw et al., 2003). Giordani and Söderlind (2003) highlight the merits of using disagreement as a measure of IU. However, they also document that disagreement only accounts for approximately one half of the variation in average individual IU from the Federal Reserve's Survey of Professional Forecasters, leaving a considerable part of the latter unexplained. Using the same data, Rich and Tracy (2010) show that disagreement is weakly correlated with IU measures derived from density forecasts. Moreover, Zarnowitz and Lambros (1987) describe several cases in which the two measures deviate. These observations are rationalized in a model of forecast error components by Lahiri and Sheng (2010), who show that disagreement can be understood as one component of average individual IU.

This study provides a theory-guided empirical examination of the covariates of IU in the Eurozone. The latter is quantified by means of data from the European Central Bank's Survey of Professional Forecasters (SPF). Employing the forecast error component model of Lahiri and Sheng (2010), we conduct a variance decomposition which isolates disagreement as one of two components of average individual IU. The additional component is the perceived variance of the shocks that occur during the period after forecasters report their predictions until the realization of the inflation rate. The model predicts that the significance of this component increases with the forecast horizon.

In the theoretical literature, relationships between several macroeconomic variables and IU have been identified. Friedman (1977) and Ball (1992) describe a positive influence of the level of inflation on IU. Empirically, the Friedman-Ball-hypothesis has been confirmed in several studies including Baillie et al. (1996), Conrad and Karanasos (2005), Conrad et al. (2010) and Hartmann and Herwartz (2012). However, Cosimano and Jansen (1988) show that the relationship might not be as strong as assumed. Baillie et al. (1996) find that the relationship is weaker for industrialized countries including the major European economies. In addition, Friedman (1977) argues that IU is lower during periods of high economic growth. Schwert (1989) finds that IU is positively related to stock market volatility. Other theoretical studies identify links between IU and policy-related variables. Taylor (1993, 2012) argues that a predictable, i.e., rules-based, monetary policy reduces economic uncertainty. According to Taylor (2012), one of the causes of the financial crisis is that the interest rate set by the Federal Reserve during the years after 2003 had been lower than the optimal rate prescribed by a Taylor rule. Conrad and Hartmann (2014) show that deteriorated macroeconomic conditions and discretionary monetary policy have a joint impact on IU.

We contribute to this literature by providing a comparison between the covariates of average individual IU and disagreement. We employ a dynamic panel model, which is estimated by a generalized method of moments (GMM) approach as discussed by Blundell and Bond (1998). After investigating whether the predictions of the model of Lahiri and Sheng (2010) hold for the SPF data, the variance decomposition serves as a means to select a set of suitable instrumental variables. Moreover, we analyze to which extent the difference between average individual IU and disagreement evolves in a predictable way, i.e., if it can be explained by distinct indicators of macroeconomic conditions and monetary policy. If this is the case, disagreement should not be considered a reliable proxy for IU.

We find that the difference between average individual IU and disagreement increases with the forecast horizon. Moreover, dynamic panel estimates show that the influence of monetary policy indicators depends on which IU measure is considered as the dependent variable. Most importantly, average individual IU rises during periods when monetary policy is more expansive than what is prescribed by a Taylor rule (Taylor, 1993). In contrast, disagreement is more closely related to contractionary monetary policy. Hence, we find that average individual IU and disagreement are related to distinct indicators of monetary policy. Further analysis shows that the difference between average individual IU and disagreement increases during periods when monetary policy is expansive, whereas disagreement itself is not affected during such episodes. This suggests that an important reason for the steady increase in average individual IU since the beginning of the financial crisis is the sustained period of expansionary monetary policy. This influence is not detected if disagreement alone is used as an indicator of IU. Thus, assessments of the role of monetary policy for IU based on disagreement are incomplete. In contrast, both measures of IU are related to macroeconomic conditions such as the inflation rate and the growth rate of real GDP in a relatively similar way that is also consistent with findings from other empirical studies such as Conrad et al. (2010).

The remainder of this chapter is structured as follows. Section 3.2 describes the IU measures that are employed in this study, whereas Section 3.3 introduces the econometric models. The data are introduced in Section 3.4. Subsequently, Section 3.5 presents and discusses the empirical results. In Section 3.6, we summarize and conclude.

3.2 Forecasting and Uncertainty

In this section, we describe the *ex-ante* measures of IU that are based on survey data. We first introduce an error component model which motivates the variance decomposition that underlies the distinction between average individual IU and disagreement. Second, based on the interpretation of the average variance, we clarify how its components can be understood. The h-step-ahead forecast errors of survey participants $i = 1, \dots, N$ are defined as

$$e_{i,t+h|t} = \pi_{t+h} - \mu_{i,t+h|t}, \tag{3.1}$$

where the inflation rate is denoted as π_{t+h} and $\mu_{i,t+h|t}$ represents h-step-ahead forecasts. The quarterly survey periods are indicated by the index $t = 1, \dots, T$. Following Lahiri and Sheng (2010), the forecast error in Eqn. (3.1) is decomposed into a part that is common to all forecasters and an idiosyncratic component, i.e.,

$$e_{i,t+h|t} = \lambda_{t+h|t} + \varepsilon_{i,t+h|t}, \tag{3.2}$$

where, firstly, $\lambda_{t+h|t} = \sum_{j=1}^{h} u_{t+j}$ denotes the sum of unpredictable disturbances that occur after forecasts are issued in t until the target date $t + h$. This component is considered as common to all forecast errors. This is justified as long as no panelist is able to predict the disturbances that occur between t and $t+h$.[1] A more formal definition of the process $\lambda_{t+h|t}$ could be obtained given assumptions about the loss function and the way information is processed by the forecasters. If predictions are evaluated by the mean squared error (MSE) criterion, the conditional expectation of π_{t+h} is the optimal predictor. In this case, $\lambda_{t+h|t}$ can be expected to follow a moving average process of order $h - 1$. However, the SPF questionnaire does not state a criterion by which forecasts should be evaluated. Moreover, a symmetric loss function like the MSE is not necessarily a realistic characterization of the preferences of all survey participants.

Secondly, individual characteristics such as differences in forecasters' methods of processing available information are incorporated in $\varepsilon_{i,t+h|t}$. Such differences might arise from forecasters' modeling choices, information constraints or their occupation. For example, both employees of research institutes and the financial industry participate in the SPF. It seems unlikely that such experts' forecasts are not affected by the largely distinct

[1] The variance decomposition does not depend on the particular functional form of the term $\lambda_{t+h|t}$. The definition $\lambda_{t+h|t} = \sum_{j=1}^{h} u_{t+j}$ highlights the interpretation of the common component as a sum of news inflows that influences the precision of all forecasts.

incentives that emerge due to their respective working environment.[2]

It is important to note that we analyze fixed-horizon forecasts, characterized by a fixed forecast horizon and a rolling target period, whereas Lahiri and Sheng (2010) discuss fixed-event forecasts, characterized by a fixed target period and a rolling forecast horizon. Lahiri and Sheng (2010) make the following assumptions: First, $\mathbf{E}[u_{t+j}] = 0$ and $\mathbf{Var}[u_{t+j}] = \sigma^2_{u,t+j}$ for any t and j. Moreover, the shocks are uncorrelated at different points in time, so that $\mathbf{E}[u_{t+j}u_{t+j'}] = 0$ for any t and $j \neq j'$ and $\mathbf{E}[u_{t+j}u_{t'+j}] = 0$ for any j and $t \neq t'$. Second, $\mathbf{E}[\varepsilon_{i,t+h|t}] = 0$ and $\mathbf{Var}[\varepsilon_{i,t+h|t}] = \sigma^2_{\varepsilon,i,t+h|t}$ for any i, t and h and the individual shocks of different forecasters are uncorrelated with each other, so that $\mathbf{E}[\varepsilon_{i,t+h|t}\varepsilon_{i',t+h|t}] = 0$ for any t, h and $i \neq i'$. Third, the individual and aggregate shocks are uncorrelated, i.e., $\mathbf{E}[\varepsilon_{i,t+h|t}u_{t'+j}] = 0$ for any i, t, t', h and j. Given these assumptions, *ex-ante* expected disagreement, i.e., the expectation of the cross-sectional variance of $\mu_{i,t+h|t}$, can be written as

$$S^2_{t+h|t} = \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{E}\left[\left(\varepsilon_{i,t+h|t} - \frac{1}{N}\sum_{j=1}^{N}\varepsilon_{j,t+h|t}\right)^2\right] = \frac{1}{N}\sum_{i=1}^{N}\sigma^2_{\varepsilon,i,t+h|t}. \tag{3.3}$$

To see that disagreement can be thought of as a part of overall uncertainty, the component model in Eqn. (3.2) serves as the basis for the variance decomposition of average individual uncertainty. The individual uncertainty about future inflation, defined as the conditional variance of the errors in Eqn. (3.2), is given by

$$\sigma^2_{i,t+h|t} = \sigma^2_{\lambda,t+h|t} + \sigma^2_{\varepsilon,i,t+h|t}, \tag{3.4}$$

where $\sigma^2_{\lambda,t+h|t} = \sum_{j=1}^{h}\sigma^2_{u,t+j}$ denotes the perceived volatility of future aggregate shocks under the assumptions stated above. Following Zarnowitz and Lambros (1987) and Lahiri and Sheng (2010), we obtain average individual IU as

$$\overline{\sigma^2_{t+h|t}} = \frac{1}{N}\sum_{i=1}^{N}\sigma^2_{i,t+h|t}. \tag{3.5}$$

This IU statistic can be interpreted as the uncertainty surrounding a forecast that is randomly drawn from the cross-section (Lahiri and Sheng, 2010). Combining Eqns. (3.3)

[2]Lahiri and Sheng (2010) argue that Eqn. (3.2) can also include a third component, $\phi_{i,h}$, which might reflect an individual horizon-specific bias. However, the arguments in the theoretical model of Lahiri and Sheng (2010) are derived by disregarding $\phi_{i,h}$, since the estimates of this component are small on average. Similarly, we have found that the averages across the estimates of $\phi_{i,h}$ for all forecasters are of relatively small size for the SPF data. Therefore, we disregard this component in the exposition of the model and only consider it in the form of individual-specific effects in the empirical panel model in Section 3.3.

to (3.5), average individual IU is decomposed such that

$$\overline{\sigma^2_{t+h|t}} = \sigma^2_{\lambda,t+h|t} + S^2_{t+h|t}. \tag{3.6}$$

In Eqn. (3.6), average individual IU is given by the sum of the variance of aggregate shocks and expected disagreement. According to the interpretation of $\overline{\sigma^2_{t+h|t}}$, disagreement represents the part of average uncertainty that results from alternative ways of processing information among individual forecasters, e.g., by employing alternative models. The extent to which $S^2_{t+h|t}$ can be expressed by a set of time series models instead of survey forecasts has been discussed, e.g., by Branch (2004). Furthermore, the variance of aggregate shocks can be interpreted as the wedge between average individual IU and disagreement. Since $\sigma^2_{\lambda,t+h|t} = \sum_{j=1}^{h} \sigma^2_{u,t+j}$, this component increases with the forecast horizon h. This is a direct consequence of the model specification. A large $\sigma^2_{\lambda,t+h|t}$ may also occur if forecasters are increasingly uncertain about future shocks, e.g., during recession periods. Thus, the model suggests that the suitability of disagreement as a measure of IU may depend on both the forecast horizon and the state of the economy.

In this study, we use survey data to estimate the unobserved quantities in Eqn. (3.6). Let $f_{i,t+h|t}$ denote a density forecast for the inflation rate, π_{t+h}, reported by individual forecaster i. Individual inflation expectations are expressed by the mean, $\mu_{i,t+h|t}$, of $f_{i,t+h|t}$. We quantify disagreement by

$$s^2_{t+h|t} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mu_{i,t+h|t} - \overline{\mu_{t+h|t}} \right)^2, \tag{3.7}$$

where $\overline{\mu_{t+h|t}} = (1/N) \sum_{i=1}^{N} \mu_{i,t+h|t}$ is the average forecast.[3] Lahiri and Sheng (2010) show that, for sufficiently large N, $s^2_{t+h|t}$ can be used to approximate $S^2_{t+h|t}$ in Eqn. (3.3). Moreover, the variance of an individual density forecast, $f_{i,t+h|t}$, delivers an estimate of individual IU, $\sigma^2_{i,t+h|t}$, from which $\overline{\sigma^2_{t+h|t}}$ is obtained using Eqn. (3.5). Finally, we obtain an estimate of the variance of aggregate shocks, $\sigma^2_{\lambda,t+h|t}$, by replacing $S^2_{t+h|t}$ in Eqn. (3.6) with $s^2_{t+h|t}$.

3.3 Modeling Inflation Uncertainty

In the following, the empirical models that are employed to analyze the covariates of IU are introduced. We first compare survey- and model-based forecasts to show that the

[3] Disagreement is often defined as the variance of point forecasts. We employ the means of the density forecasts instead. Clements (2014) provides a critical discussion.

SPF dataset provides relatively accurate forecasts as compared to commonly used time series models and also to illustrate to what extent disagreement may arise from distinct models. Second, we introduce an empirical framework to examine how average individual IU, disagreement and the variance of aggregated shocks are related to each other and how they can be explained by various indicators of macroeconomic conditions and monetary policy that have been identified as drivers of IU in the related literature.

3.3.1 Comparing Survey Forecasts to Model-Based Predictions

The partitioning of the forecast errors in Eqn. (3.2) is based on the view that survey forecasts can be characterized by both common and individual features. The idiosyncratic component may arise both from distinctions in the information sets up to and prior to period t and different ways of incorporating such information in a model framework. In particular, the extent to which the individual part can be regarded as a result of alternative model choices can be examined by comparing a number of forecasts that are derived from commonly used econometric models to the survey-based predictions.[4] This approach has been employed, e.g., in Branch (2004), who compares survey forecasts to the predictions obtained from three alternative reduced form forecasting models.

A model that might be especially suitable since it accounts for changes in the conditional variance of inflation is the unobserved components stochastic volatility (UCSV) model of Stock and Watson (2007). In a study on inflation data for the U.S., Stock and Watson (2007) report that the UCSV delivers more accurate inflation forecasts than several competing time series models. The UCSV decomposes π_{t+1} into a long-term trend τ_{t+1} and a transitory component w_{t+1}, i.e.,

$$\begin{aligned} \pi_{t+1} &= \tau_{t+1} + w_{t+1}, \quad & w_{t+1} &= \mathfrak{h}_{w,t+1} Z_{w,t+1}, \\ \tau_{t+1} &= \tau_t + \epsilon_{t+1}, & \epsilon_{t+1} &= \mathfrak{h}_{\epsilon,t+1} Z_{\epsilon,t+1}. \end{aligned} \tag{3.8}$$

In this framework, the conditional variances follow independent random walk processes, i.e., $\ln \mathfrak{h}^2_{w,t+1} = \ln \mathfrak{h}^2_{w,t} + \nu_{w,t+1}$ and $\ln \mathfrak{h}^2_{\epsilon,t+1} = \ln \mathfrak{h}^2_{\epsilon,t} + \nu_{\epsilon,t+1}$. The variances $\mathfrak{h}^2_{w,t+1}$ and $\mathfrak{h}^2_{\epsilon,t+1}$ denote the transitory and permanent components of inflation fluctuations, respectively. We assume $Z_{t+1} = (Z_{w,t+1}, Z_{\epsilon,t+1}) \overset{i.i.d.}{\sim} N(\mathbf{0}, \boldsymbol{I}_2)$ and $\nu_{t+1} = (\nu_{w,t+1}, \nu_{\epsilon,t+1}) \overset{i.i.d.}{\sim} N(\mathbf{0}, \theta \boldsymbol{I}_2)$, where the coefficient θ controls the smoothness of both variance processes. We further assume that the disturbances Z_{t+1} and ν_{t+1} are mutually independent. Similar to Stock and Watson (2007), one-year-ahead inflation forecasts from the UCSV are obtained as

[4] We thank Richard Baillie for the suggestion to include model-based forecasts as a benchmark to evaluate the survey-based predictions.

$\mu^{ucsv}_{t+4|t} = \tau_{t|t}$ after calibrating the parameter θ to a value of $(0.2/3)^2$ and estimating the state-space model in Eqn. (3.8) by means of a Gibbs sampling recursion.

Following Branch (2004), we obtain a second set of forecasts from first-order vector autoregressive (VAR) models. The VAR is given by

$$\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{z}_{t-1} + \boldsymbol{\vartheta}_t, \qquad \boldsymbol{\vartheta} \sim (\boldsymbol{0}, \varsigma^2 \boldsymbol{I}_2), \tag{3.9}$$

where $\boldsymbol{z}_t$ denotes a 2×1 vector of variables including the inflation rate and a further variable as will be described in Section 3.5.

Using pre-sample values and recursively reestimating the models in Eqns. (3.8) and (3.9) yields pseudo-out of sample inflation forecasts, $\mu^{var}_{t+4|t}$ and $\mu^{ucsv}_{t+4|t}$, from which we compute the corresponding one-year prediction errors, as denoted $e^{ucsv}_{t+4|t}$ and $e^{var}_{t+4|t}$. These errors can be compared to the corresponding average one-year-ahead forecast error, $\overline{e_{t+4|t}} = \pi_{t+4} - \overline{\mu_{t+4|t}}$, from the SPF.

3.3.2 The Importance of the Forecast Horizon and Recessions

After establishing the informative content of the SPF forecasts, we turn to the analysis of survey-based IU measures. The decomposition in Eqn. (3.6) states that the variance of aggregate shocks equals the difference between average individual IU and disagreement. The model assumptions imply that $\sigma^2_{\lambda,t+h|t}$ increases with h. Moreover, it may also increase during recession periods. In order to evaluate empirically how much $\overline{\sigma^2_{t+h|t}}$ and $s^2_{t+h|t}$ deviate from each other, we estimate the following pooled model which includes observations for all available forecast horizons $h = 1, \ldots, H$:

$$\sigma^2_{\lambda,t+h|t} = \zeta_1 D^{h=1} + \ldots + \zeta_H D^{h=H} + \gamma_1(t-1) + \gamma_2 D^{REC}_{t-1} + \nu_{t+h|t}, \tag{3.10}$$

where $\nu_{t+h|t} \sim (0, \sigma^2_{\nu,h})$ is the error term. The indicator variables $D^{h=1}, \ldots, D^{h=H}$ are defined such that, for example, $D^{h=1}$ is equal to unity for $h = 1$, and zero for all other horizons. Since the model of Lahiri and Sheng (2010) suggests that the variance of aggregate shocks increases with the forecast horizon, it is expected that $\zeta_H > \ldots > \zeta_1 > 0$. To account for potential non-stationarity of $\sigma^2_{\lambda,t+h|t}$, a time trend, $t-1$, is included. The impact of recessions is captured by the indicator variable D^{REC}_{t-1}, which is equal to unity during recession periods, and zero in all other cases. Recessions are identified by the method of Bry and Boschan (1971) as well as Harding and Pagan (2002) using the output gap, $\tilde{x}_{t-1} = 100 \times \left(x_{t-1} - x^{HP}_{t-1}\right)$, as an indicator for the business cycle. In this equation, $x_{t-1} = \ln(X_{t-1})$ denotes the natural logarithm of the level of real GDP,

X_{t-1}, and x_{t-1}^{HP} is the trend component of x_{t-1}, which is extracted by using the Hodrick-Prescott filter. Thereby, we identify the peaks and troughs of economic activity in the Eurozone. Following Chauvet and Hamilton (2005), a recession is defined as the period between a peak and a trough (excluding the former but including the latter).[5] If the variance of aggregate shocks increases during recession periods, we would expect that $\gamma_2 > 0$.

3.3.3 Macroeconomic and Policy Influences on IU

The decomposition of Lahiri and Sheng (2010) implies that average individual IU and disagreement should be considered as distinct measures of IU. Thus, we evaluate them separately below. In order to control for forecaster-specific characteristics, we exploit the information from the panel of individual forecasters. The individual IU series from the SPF exhibit a clearly visible upward trend which is in many cases similar to the one depicted for the average IU in Figure 3.2. Long memory in the conditional second moments of inflation for several industrialized economies has been documented by Baillie et al. (2002). Thus, we employ a dynamic panel model with an autoregressive term and fixed effects. However, owing to the recurring dropout and replacement of survey participants, the number of time periods per forecaster might be regarded as eventually limited. For fixed T, a bias in the coefficient estimates of a dynamic panel model is likely to arise due to the correlation between the lagged dependent variable and the individual effects (Blundell and Bond, 1998). Therefore, the estimation of the model parameters proceeds by means of a GMM approach, following Blundell and Bond (1998). The dynamic panel specification reads as

$$y_{i,t+h|t} = \rho_h y_{i,t+h-1|t-1} + \boldsymbol{\alpha}_h' \boldsymbol{m}_{t-1} + \boldsymbol{\beta}_h' \boldsymbol{p}_{t-1} + \boldsymbol{\gamma}_h' \boldsymbol{d}_{t-1} + v_{i,t+h|t} \tag{3.11}$$

for $i = 1, \ldots, N$, where $y_{i,t+h|t} \in \{\sigma^2_{i,t+h|t}, \sigma^2_{\varepsilon,i,t+h|t}\}$ denotes either individual IU or the variance of individual shocks. The latter is obtained as $\sigma^2_{\varepsilon,i,t+h|t} = \sigma^2_{i,t+h|t} - \sigma^2_{\lambda,t+h|t}$ based on Eqn. (3.4) and (3.6). Morevoer, $\boldsymbol{m}_{t-1}$ is a 3×1 vector of macroeconomic conditions, $\boldsymbol{p}_{t-1}$ is a 6×1 vector containing indicators of monetary policy, $\boldsymbol{d}_{t-1} = (1, t-1, D^{REC}_{t-1})'$ is a 3×1 vector of exogenous variables, $\boldsymbol{\alpha}_h$, $\boldsymbol{\beta}_h$ and $\boldsymbol{\gamma}_h$ denote the corresponding coefficient vectors and $v_{i,t+h|t} = \eta_{i,h} + \nu_{i,t+h|t}$ is the error term. The forecaster-specific fixed effects, $\eta_{i,h}$,

[5] We have also used an alternative approach in which a recession is defined as two (or more) consecutive periods characterized by a negative real GDP growth rate. The results in this chapter are robust to the employed definition of an economic recession. We thank an anonymous referee from the *Journal of Empirical Finance* for this suggestion.

capture unobserved heterogeneity with respect to individual IU.[6] We estimate Eqn. (3.11) using the approach of Blundell and Bond (1998) and include up to six-period lagged levels as well as first differences of $y_{i,t+h-1|t-1}$ and the variables in $\boldsymbol{m}_{t-1}$ and $\boldsymbol{p}_{t-1}$ as instrumental variables. In addition, we include $\sigma^2_{\lambda,t+h-1|t-1}$, which is related to both $\sigma^2_{i,t+h-1|t-1}$ and $\sigma^2_{\varepsilon,i,t+h-1|t-1}$ according to the definition in Eqn. (3.4). Hence, $\sigma^2_{\lambda,t+h-1|t-1}$ is likely to be a relevant instrument for $y_{i,t+h-1|t-1}$. Moreover, past realizations of $\overline{\sigma}^2_{t+1h-1|t-1}$ and $s^2_{t+h-1|t-1}$ are readily available to forecasters when they report their predictions in period t and therefore can be considered as part of their information sets. Based on Eqn. (3.6), $\sigma^2_{\lambda,t+h-1|t-1}$ should thus also be exogeneous. To account for forecaster-specific heteroskedasticity and autocorrelation in $\nu_{i,t+h|t}$, we use robust standard errors and apply the adjustment by Windmeijer (2005). Unlike Eqn. (3.10), the model in Eqn. (3.11) is estimated separately for each forecast horizon h. For notational convenience, we do not differentiate between the coefficients for alternative choices of $y_{i,t+h|t}$.

To explain fluctuations in IU, Friedman (1977) and Schwert (1989) discuss the role of distinct macroeconomic conditions and the dynamics on financial markets. Such influences on IU are summarized in

$$\boldsymbol{m}_{t-1} = (\pi_{t-1}, \Delta x_{t-1}, RV_{t-1}(E))'. \tag{3.12}$$

According to Friedman (1977), two important covariates are the level of inflation, which is defined as the year-on-year change of the quarterly Harmonised Index of Consumer Prices (HICP),

$$\pi_{t-1} = 100 \times \frac{HICP_{t-1} - HICP_{t-5}}{HICP_{t-5}}, \tag{3.13}$$

and output growth,

$$\Delta x_{t-1} = 100 \times \frac{X_{t-1} - X_{t-5}}{X_{t-5}}. \tag{3.14}$$

Empirical evidence regarding these relations has been documented, e.g., by Dovern et al. (2012) or Hartmann and Roestel (2013). Moreover, Engle and Rangel (2008) and Conrad and Loch (2015) document for the U.S. that IU is related to stock market volatility. We measure stock market fluctuations by employing the intra-quarter variation of squared

[6] The unobserved heterogeneity includes, for example, the horizon-specific individual bias, $\phi_{i,h}$.

returns,

$$RV_{t-1}(E) = 100 \times \sqrt{\sum_{\ell \in t-1} r^2_{\ell,t-1}}, \tag{3.15}$$

with $r_{\ell,t-1} = \ln(P_{\ell,t-1}/P_{\ell-1,t-1})$ denoting the daily return of equity prices, $P_{\ell,t-1}$.

Apart from macroeconomic conditions, IU might also be influenced by monetary policy. This hypothesis has been described, for example, in the widely-cited study of Ball (1992). Distinct policy indicators are summarized in

$$\boldsymbol{p}_{t-1} = (\Delta EPU_{t-1}, \Delta AS_{t-1}, TD^+_{t-1}, |TD^-_{t-1}|, MPC_{t-1}, \Delta MM_{t-1})'. \tag{3.16}$$

IU might be affected by spillovers from uncertainty about economic and political conditions in general. To account for such influences, we consider the changes in the economic policy uncertainty indicator as denoted ΔEPU_{t-1} (cf. Baker et al., 2016). A further source of IU might be the unconventional monetary policy measures adopted by the ECB such as the changes in the asset position of its balance sheet. The relation between IU and balance sheet adjustments, denoted ΔAS_{t-1}, might capture perceived increases in inflation risks due to the ECB's large-scale asset purchases since 2008.[7] Moreover, a relation between deviations from a rules-based monetary policy and uncertainty in general has been put forth by Taylor (2012). The degree to which the ECB might have deviated from a predictable monetary policy scheme such as the Taylor (1993) rule is expressed by the variables TD^+_{t-1} and $|TD^-_{t-1}|$ which measure contractionary and expansionary monetary policy, respectively. Deviations from the Taylor rule are defined as

$$TD_{t-1} = i_{t-1} - i^*_{t-1}, \tag{3.17}$$

where i_{t-1} is the interest rate set by the central bank and i^*_{t-1} is the optimal predicted Taylor rule interest rate specified as a function of the inflation rate and the output gap in line with the dynamic model proposed by Clarida et al. (1998),

$$i_{t-1} = \omega_0 + \omega_1 \pi_{t+3} + \omega_2 \tilde{x}_{t-1} + \xi_{t-1}. \tag{3.18}$$

Since the regressors in Eqn. (3.18) are likely endogenous, the model is estimated by means of GMM using up to four-period lags of i_{t-1}, π_{t-1} and $\tilde{x}_{t-1}$ as instrumental variables. Based on Eqns. (3.17) and (3.18), we define $TD^+_{t-1} = TD_{t-1} \times \mathbb{1}(TD_{t-1} > 0)$, where $\mathbb{1}(TD_{t-1} >$

[7] We thank an anonymous referee from the *Journal of Empirical Finance* for the suggestion to replace ΔAS_{t-1} with the changes of the ECB assets-to-nominal-GDP ratio. Our results are robust to the choice of the employed asset variable.

0) equals unity if $TD_{t-1} > 0$, and zero else. Similarly, $|TD^{-}_{t-1}| = |TD_{t-1}| \times \mathbb{1}(TD_{t-1} < 0)$, where $\mathbb{1}(TD_{t-1} < 0)$ equals unity if $TD_{t-1} < 0$, and zero else.[8] The assessment of monetary policy by means of the Taylor rule is difficult once the interest rate targeted by the central bank approaches a value of zero. However, i_{t-1} has not been exactly equal to zero during any quarter of the sample period. In addition, the set of variables in $\boldsymbol{p}_{t-1}$ includes ΔAS_{t-1} and further variables related to monetary aggregates and central bank communication. We regard these covariates as complementary to TD_{t-1}, especially during periods when interest rates are low.

In order to account for the potential influence of the communication of the ECB, we consider the Monetary Policy Communicator, denoted as MPC_{t-1}, that quantifies the ECB's communication of risks regarding future price stability (Conrad and Lamla, 2010). Finally, IU might increase as a result of large adjustments in monetary aggregates (Holland, 1995). To acknowledge this potential transmission channel, we consider changes of the money multiplier, ΔMM_{t-1}, where

$$MM_{t-1} = \frac{M3_{t-1}}{M0_{t-1}}. \tag{3.19}$$

In Eqn. (3.19), $M3_{t-1}$ is an indicator of the broad money supply and base money and $M0_{t-1}$ denotes base money. To summarize, $\boldsymbol{p}_{t-1}$ comprises indicators of the monetary policy stance that are based on interest rates but also takes more traditional, money-based indicators into account.

As mentioned in Section 3.2, average individual IU can be interpreted as the uncertainty of the average forecaster. Since it holds that $\overline{\sigma^2_{t+h|t}} = (1/N)\sum_{i=1}^{N} \sigma^2_{i,t+h|t}$ and $s^2_{t+h|t} = (1/N)\sum_{i=1}^{N} \sigma^2_{\varepsilon,i,t+h|t}$, Eqn. (3.11) implies that

$$y_{t+h|t} = \rho_h y_{t+h-1|t-1} + \boldsymbol{\alpha}'_h \boldsymbol{m}_{t-1} + \boldsymbol{\beta}'_h \boldsymbol{p}_{t-1} + \boldsymbol{\gamma}'_h \boldsymbol{d}_{t-1} + \overline{v_{t+h|t}}, \tag{3.20}$$

where $y_{t+h|t} \in \{\overline{\sigma^2_{t+h|t}}, s^2_{t+h|t}\}$ and $\overline{v_{t+h|t}} = \overline{\eta_h} + \overline{\nu_{t+h|t}}$ with $\overline{\eta_h} = (1/N)\sum_{i=1}^{N} \eta_{i,h}$ and $\overline{\nu_{t+h|t}} = (1/N)\sum_{i=1}^{N} \nu_{i,t+h|t}$. Thus, the effect of a variable in $\boldsymbol{m}_{t-1}$ or $\boldsymbol{p}_{t-1}$ on $\sigma^2_{i,t+h|t}$ ($\sigma^2_{\varepsilon,i,t+h|t}$) is theoretically equivalent to its impact on $\overline{\sigma^2_{t+h|t}}$ ($s^2_{t+h|t}$). Consequently, the estimates of Eqn. (3.11) can be interpreted as the effect of macroeconomic conditions and monetary policy on average individual IU and disagreement.

[8] It could be argued that a two-year horizon in Eqn. (3.18) may be more appropriate to characterize the inflation targeting of the ECB. As a robustness check, we replaced π_{t+3} with π_{t+7} and used the estimates to construct alternative versions of TD^{+}_{t-1} and $|TD^{-}_{t-1}|$. However, the alternative time series were very similar and all empirical results in this chapter are robust to the chosen specification of Eqn. (3.18). We thank an anonymous referee from the *Journal of Empirical Finance* for this suggestion.

3.3.4 Explaining the Variance of Aggregate Shocks

Based on the decomposition in Eqn. (3.4), the impact of the macroeconomic conditions and policy indicators on individual IU is given by their aggregate impact on the variances of the aggregate and individual shocks. Thus, if the impact of the covariates on $\sigma^2_{i,t+h|t}$ and $\sigma^2_{\varepsilon,i,t+h|t}$ differs, this must be due to a systematic influence of $\boldsymbol{m}_{t-1}$ or $\boldsymbol{p}_{t-1}$ (or both) on the second component, i.e., $\sigma^2_{\lambda,t+h|t}$. Similarly, the impact of a covariate on average individual IU is decomposed into the sum of the effects on the variance of aggregate shocks and disagreement based on Eqn. (3.6). Thus, we also analyze the covariates of the variance of aggregate shocks by means of

$$\sigma^2_{\lambda,t+h|t} = \tilde{\boldsymbol{\alpha}}'_h \boldsymbol{m}_{t-1} + \tilde{\boldsymbol{\beta}}'_h \boldsymbol{p}_{t-1} + \tilde{\boldsymbol{\gamma}}'_h \boldsymbol{d}_{t-1} + \nu_{t+h|t}, \tag{3.21}$$

where $\nu_{t+h|t}$ is the horizon-specific error term. If, for example, expansionary monetary policy, as measured by $|TD^-_{t-1}|$, has no effect on $s^2_{t+h|t}$ but increases $\sigma^2_{\lambda,t+h|t}$, the net effect of such a policy on $\overline{\sigma^2_{t+h|t}}$ may be positive. Considering only disagreement as a proxy for overall IU, however, one would mistakingly conclude that expansionary monetary policy does not increase IU in this scenario. Thus, it is recommendable to test whether there are observable factors which increase the difference between $\overline{\sigma^2_{t+h|t}}$ and $s^2_{t+h|t}$ in a systematic way. If this is the case, disagreement does not capture all relevant influences on IU and should thus not be considered as a reliable proxy for IU.

3.4 Data

Forecast data is provided by the Survey of Professional Forecasters, which has been conducted by the European Central Bank during successive quarters since 1999Q1. We employ density forecasts, $f_{i,t+h|t}$, regarding future HICP inflation in the Eurozone, π_{t+h}. As can be seen in Table 3.1, the intervals employed in the SPF questionnaire have changed on multiple occasions.

The sample contains fixed-horizon forecasts from the surveys conducted during the period 1999Q1–2013Q2, such that $T = 58$. Forecast horizons $h \in \{4, 8, 20\}$ denote one-year-, two-years- and five-years-ahead, respectively. The time series for $h = 20$ is only available in consecutive quarters since 2001Q1. We exclude inappropriate density forecasts from our analysis, e.g., if reported subjective probabilities do not sum to one. The remaining cross-section comprises predictions from $N = 98$ anonymous forecasters. Figure 3.1 depicts the participation of these forecasters in each survey round for different forecast horizons. Evidently, the panel is unbalanced. However, the number of missing

Table 3.1: Intervals used in the ECB-SPF questionnaire

$f_{i,t+h\|t}$									
1999Q1-2000Q3			$(-\infty;-0.1]$	...	$[0.5(k-1);0.5k-0.1]$,	$k=1,...,7$	...	$[3.5;+\infty)$	
2000Q4			$(-\infty;-0.1]$	...	$[0.5(k-1);0.5k-0.1]$,	$k=1,...,8$	...		$[4.0;+\infty)$
2001Q1-2008Q2			$(-\infty;-0.1]$	...	$[0.5(k-1);0.5k-0.1]$,	$k=1,...,7$	...	$[3.5;+\infty)$	
2008Q3-2009Q1			$(-\infty;-0.1]$	...	$[0.5(k-1);0.5k-0.1]$,	$k=1,...,8$	...		$[4.0;+\infty)$
2009Q2-2009Q4	$(-\infty;-2.1]$			...	$[0.5(k-1);0.5k-0.1]$,	$k=-3,...,8$	...		$[4.0;+\infty)$
2010Q1-end of sample		$(-\infty;-1.1]$		...	$[0.5(k-1);0.5k-0.1]$,	$k=-1,...,8$	...		$[4.0;+\infty)$

Notes: This table illustrates changes to the range of SPF density forecasts, $f_{i,t+h|t}$, in the questionnaire of the ECB. Each row depicts the upper and lower intervals whereas k indicates the intermediate intervals. Density forecasts are issued by forecasters $i = 1,\dots,98$ in sample period $t = 1,\dots,58$, representing time instances between 1999Q1 and 2013Q2, with forecast horizon $h \in \{4, 8, 20\}$.

observations is lower than it is for the Federal Reserve's SPF, for which Lahiri et al. (2017) document a substantial fraction of missing observations.

We follow Engelberg et al. (2009) and fit a generalized beta distribution to the individual histograms if forecasters attach non-zero probabilities to at least three different intervals.[9]

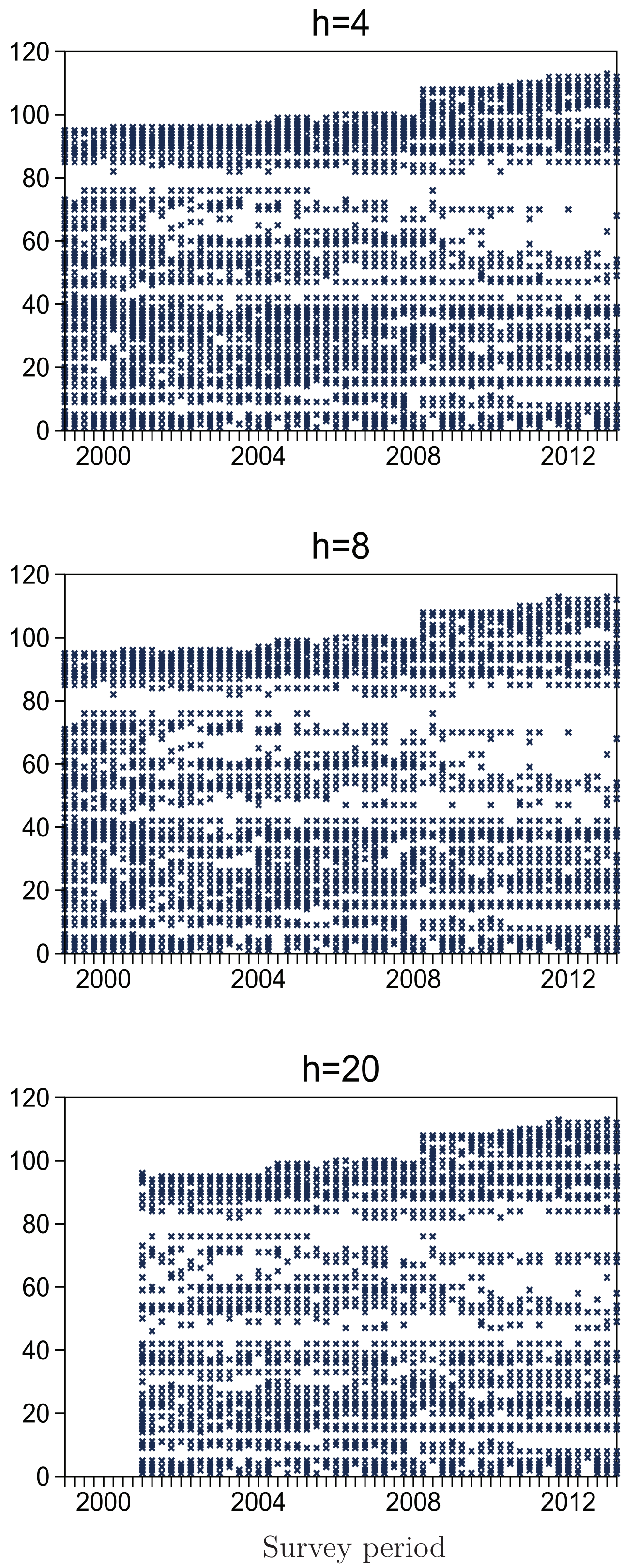
Figure 3.2 depicts individual and average inflation expectations as well as aggregate IU measures over the survey period t for $h \in \{4, 8, 20\}$. The graphs for $\mu_{i,t+h|t}$ show that long-term expectations are scattered around the ECB's inflation target of approximately 2%. For all h, the $\overline{\sigma^2_{t+h|t}}$ measures remain fairly constant between 1999 and 2007 but show an upward trend from 2007 until the end of the sample period. The $s^2_{t+h|t}$ statistics show considerable increases during the years 2008 and 2009. However, for all h, $s^2_{t+h|t}$ reverts to its pre-crisis level after a short period of time. This is in line with the argument of Lahiri and Sheng (2010) that the trajectory of $s^2_{t+h|t}$ may differ from the one of $\overline{\sigma^2_{t+h|t}}$, depending on the perceived variation of forthcoming aggregate shocks, $\sigma^2_{\lambda,t+h|t}$. The plots also show that, on average, both $\overline{\sigma^2_{t+h|t}}$ and $\sigma^2_{\lambda,t+h|t}$ increase with h, whereas there are little deviations between the $s^2_{t+h|t}$ statistics, except for the peak in 2009, which is smaller for large h.

Table 3.2 shows that the strength of the relationship between $\overline{\sigma^2_{t+h|t}}$ and $s^2_{t+h|t}$ declines from a correlation coefficient of 0.5 to less than 0.4 for increasing h, whereas the correlation statistic between $\overline{\sigma^2_{t+h|t}}$ and $\sigma^2_{\lambda,t+h|t}$ increases from 0.7 to more than 0.9. The corresponding entries in Table 3.2 are marked in boldface.

Data on inflation rates, π_t, and real GDP growth rates, Δx_t, are drawn from the Statistical Data Warehouse (SDW) of the ECB. This data source provides real-time data, which are most closely related to the information available to forecasters at the time when predic-

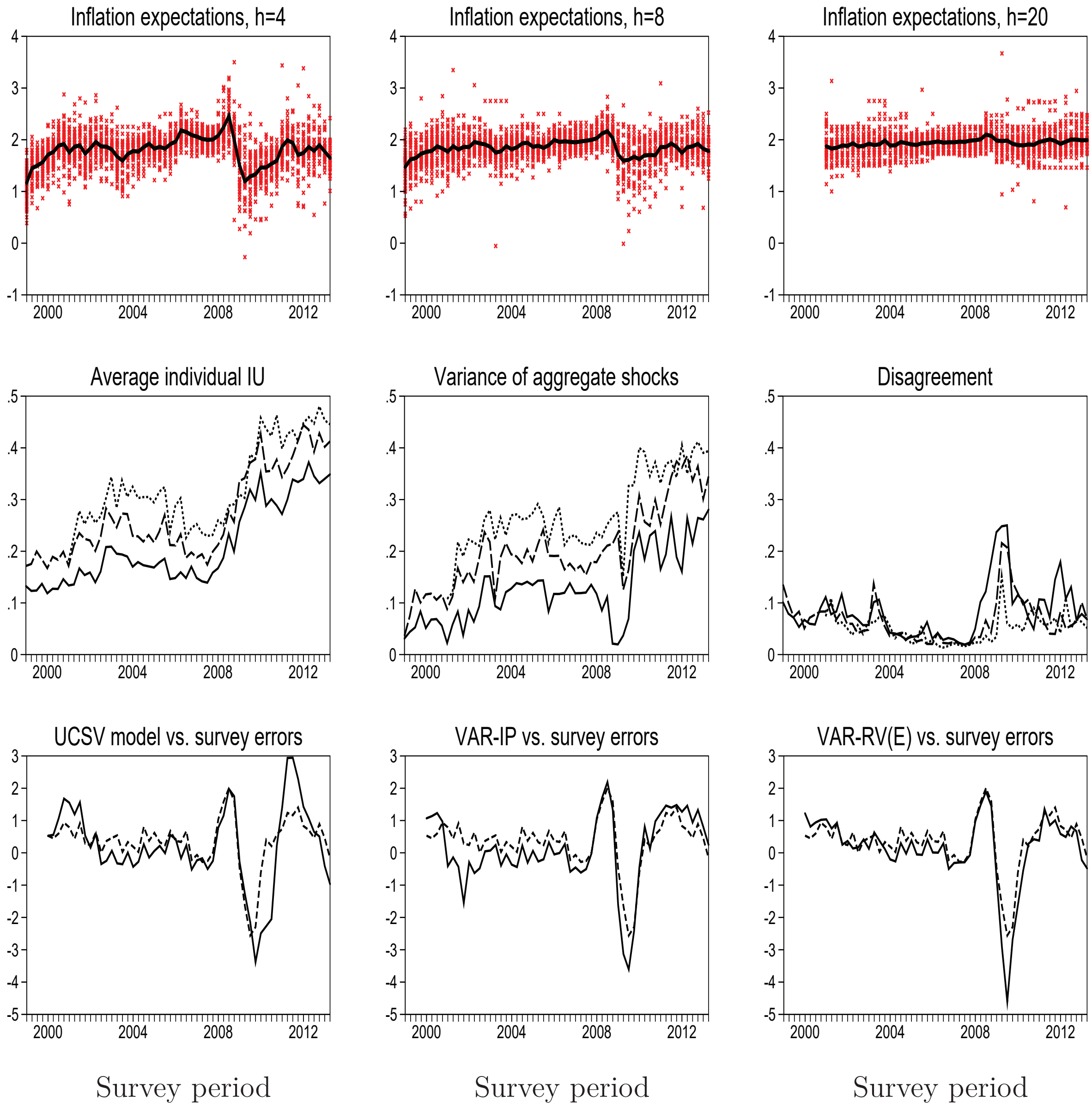
[9] Triangular distributions are fitted when less than three intervals are used. For details see Engelberg et al. (2009).

Figure 3.1: ECB-SPF forecaster panel



Notes: Graphs depict forecaster participation in the SPF between 1999Q1 and 2013Q2 for forecast horizons $h \in \{4, 8, 20\}$. The vertical axis depicts the identification numbers of the anonymous survey participants. The horizontal axis denotes the periods during which predictions are reported. Each cross indicates that a density forecast has been reported by a survey participant.

Figure 3.2: Inflation expectations and uncertainty measures



Notes: The graphs in the first row depict individual inflation expectations from the SPF, $\mu_{i,t+h|t}$, for forecast horizons $h \in \{4, 8, 20\}$ from left to right. The solid black lines are the cross-sectional means of the individual expectations, $\overline{\mu_{t+h|t}}$. The horizontal axis denotes the periods during which predictions are reported. The graphs in the second row depict average individual IU, $\overline{\sigma^2_{t+h|t}}$, and its two components, i.e., the variance of aggregate shocks, $\sigma^2_{\lambda,t+h|t}$, and disagreement, $s^2_{t+h|t}$, from left to right. Solid, dashed and dotted lines are used to distinguish forecast horizons $h \in \{4, 8, 20\}$, respectively. The graphs in the third row show one-year-ahead forecast errors derived from distinct time series models, $e^{ucsv}_{t+4|t}$ and $e^{var}_{t+4|t}$ (solid lines), against the average forecast error from the SPF, $\overline{e_{t+4|t}}$ (dashed line).

Table 3.2: Correlations across aggregate IU measures

	$\overline{\sigma^2_{t+4\|t}}$	$\overline{\sigma^2_{t+8\|t}}$	$\overline{\sigma^2_{t+20\|t}}$	$s^2_{t+4\|t}$	$s^2_{t+8\|t}$	$s^2_{t+20\|t}$	$\sigma^2_{\lambda,t+4\|t}$	$\sigma^2_{\lambda,t+8\|t}$	$\sigma^2_{\lambda,t+20\|t}$
$\overline{\sigma^2_{t+4\|t}}$	1.00								
$\overline{\sigma^2_{t+8\|t}}$	0.98	1.00							
$\overline{\sigma^2_{t+20\|t}}$	0.92	0.94	1.00						
$s^2_{t+4\|t}$	**0.50**	0.50	0.28	1.00					
$s^2_{t+8\|t}$	0.47	**0.47**	0.32	0.73	1.00				
$s^2_{t+20\|t}$	0.44	0.47	**0.38**	0.43	0.63	1.00			
$\sigma^2_{\lambda,t+4\|t}$	**0.70**	0.68	0.80	−0.26	−0.08	0.14	1.00		
$\sigma^2_{\lambda,t+8\|t}$	0.84	**0.85**	0.88	0.14	−0.06	0.16	0.82	1.00	
$\sigma^2_{\lambda,t+20\|t}$	0.84	0.84	**0.94**	0.14	0.11	0.05	0.82	0.89	1.00

Notes: The IU measures $\overline{\sigma^2_{t+h|t}}$, $s^2_{t+h|t}$ and $\sigma^2_{\lambda,t+h|t}$ refer to average individual IU, disagreement and the variance of aggregate shocks from Eqns. (3.5), (3.7) and (3.6) at forecast horizons $h \in \{4, 8, 20\}$, respectively. Correlations which highlight the relationship between average individual IU and its components for different forecast horizons are marked in boldface. The sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

tions are reported. Using the GDP data, we calculate the output gap, $\tilde{x}_t$, from which three periods of economic recessions are identified for our sample period: 2001Q1 to 2005Q3, 2008Q2 to 2009Q2 and 2011Q4 to 2013Q2. Quarterly realized stock market volatility, $RV_t(E)$, is calculated using Eurostoxx bluechip net returns for the Eurozone.[10] To measure changes in the quarterly economic policy uncertainty, ΔEPU_t, we use the monthly EU Economic Policy Uncertainty Index of Baker et al. (2016), which is defined as the weighted sum of three components: Newspaper coverage of policy-related economic uncertainty with a weight of 0.5, forecaster disagreement about federal government budget balances with a weight of 0.25 and inflation disagreement with a weight of 0.25. The newspaper component is defined as the normalized coverage of policy-related economic uncertainty by ten influential newspapers from major EU countries and measured by the number of newspapers containing at least one term from each of the three sets 'uncertain/uncertainty', 'economic/economy' and 'policy/tax/spending/regulation/central bank/budget deficit'. The index ΔEPU_t as we use it does not include inflation disagreement. We remove this component and rescale the remaining index by dividing it by 100. Changes in the ECB's balance sheet, ΔAS_t, are measured using data on the quarterly total assets/liabilities in the Euro area in trillions of Euros. The data series is drawn from the SDW, which also

[10] As a robustness check, we replaced Eurozone stock market volatility with the volatility on global equity markets using data of the MSCI world index. The time series are very similar and our main findings are not affected by the choice of the employed volatility measure.

provides data on interest rates used in the construction of TD_t^+ and $|TD_t^-|$. To quantify ECB communication regarding forthcoming threats to price stability, we include MPC_t, defined as the quarterly average of the monthly Monetary Policy Communicator, which is published by the KOF (cf. Conrad and Lamla, 2010). This indicator translates the ECB president's statements on price stability during the monthly ECB press conferences into numerical values, such that $MPC_t \in [-1, +1]$ with positive (negative) values indicating upside (downside) risks to price stability. Measures of money supply used in the construction of ΔMM_t are drawn from the SDW, which also provides monthly data on industrial production in the Eurozone. We use the quarterly average of this time series, denoted as IP_t, in the estimation of the VAR(1) models. Figure 3.3 plots the macroeconomic variables and indicators of monetary policy.

The graphs show that the levels of EPU_t, AS_t and MM_t might be non-stationary. Hence, we consider their first differences, $\Delta EPU_t, \Delta AS_t$ and ΔMM_t, in the empirical models instead of their levels.[11]

3.5 Results

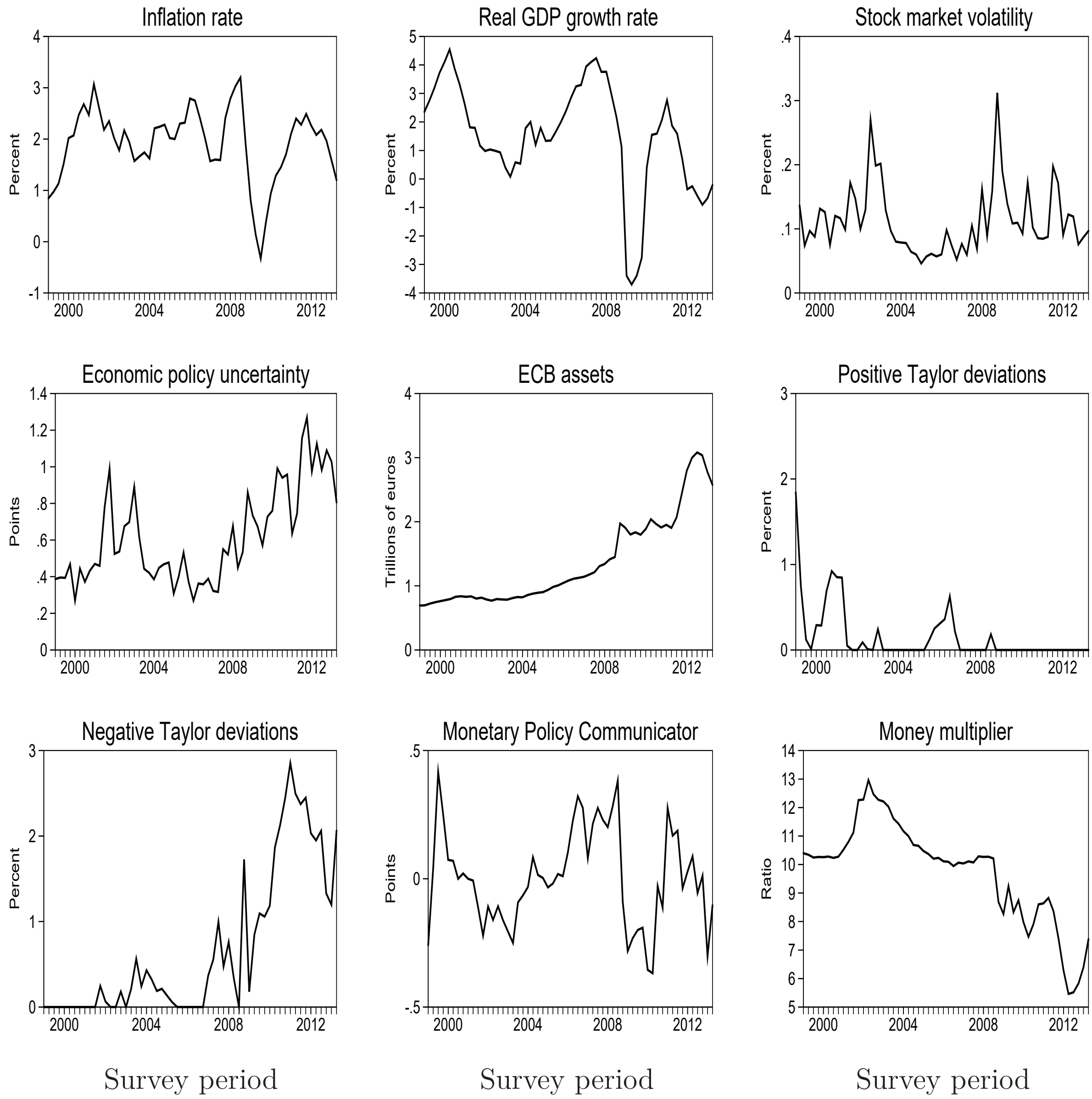
In this section, empirical results for the models introduced in Section 3.3 are summarized and discussed. We first compare forecast errors from the SPF to those derived from econometric models. Second, we examine how the variance of aggregate shocks varies with the forecast horizon as a means to test the claim that it can be interpreted as the perceived variation of successive shocks between the forecast and target period. Next, we analyze the covariates of IU by means of a dynamic panel model, followed by an investigation of potential factors that may affect the difference between average individual IU and disagreement.

3.5.1 Comparing Survey Forecasts and Model-Based Predictions

Since the IU measures we consider are derived from inflation forecasts, we compare the accuracy of survey predictions to the ones from commonly used time series models. This illustrates how the survey predictions differ from those of widely used time series approaches and, moreover, shows to which degree the idiosyncratic component of IU may reflect distinct ways of information processing of individual survey participants. As described in Section 3.3, we employ the UCSV model of Stock and Watson (2007) and two bivariate VAR(1) models with $\boldsymbol{z}_t = (\pi_t, IP_t)'$ and $\boldsymbol{z}_t = (\pi_t, RV_t(E))'$. Table 3.3 reports

[11] In addition to the graphical evidence in Figure 3.3, the (unreported) results from augmented Dickey-Fuller (ADF) tests suggested non-stationarity in most cases.

Figure 3.3: Macroeconomic and policy variables



Notes: Graphs depict variables measuring macroeconomic conditions and indicators of monetary policy.

the mean absolute errors (MAE) and mean squared errors (MSE) based on the one-year-ahead average SPF forecast, $\overline{\mu_{t+4|t}}$, and the model-based predictions. Asterisks indicate significant differences in the predictive accuracy of the SPF and model-based forecasts at the 10% level as assessed by a Diebold-Mariano test.

Table 3.3: Forecast evaluation

	SPF	UCSV	VAR-IP	VAR-RV(E)
MAE	0.67	0.92*	0.83*	0.73
MSE	0.76	1.58*	1.28*	1.19

Notes: This table reports the mean absolute errors (MAE) and mean squared errors (MSE) based on the one-year-ahead average SPF forecast, $\overline{\mu_{t+4|t}}$, and model-based predictions from an unobserved components stochastic volatility (UCSV) model and two bivariate first-order vector autoregressive (VAR) models including the inflation rate and either industrial production or realized volatility. Asterisks ('*') indicate significant differences in the predictive accuracy of the SPF and model-based forecasts at the 10% level as assessed by a Diebold-Mariano test. The sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

As can be seen from Table 3.3, the average across all survey forecasts provides lower MAE and MSE statistics than both the UCSV model and the VAR models. This underscores that the SPF data may provide a well calibrated account of existing forecasting risks. The forecast errors $\overline{e_{t+4|t}}$, $e^{ucsv}_{t+4|t}$ and $e^{var}_{t+4|t}$ are depicted in the bottom panel of Figure 3.2. Before the financial crisis, all errors are close to zero and evolve in a stable way, i.e., SPF- and model forecasts have been aligned to some degree. However, since the beginning of the crisis, the SPF error fluctuates less than the other series. Thus, the survey forecasts are more accurate than the econometric models during this period. One might speculate that forecasters revised their modeling choices more strongly in this environment, which may explain the increase in disagreement for $h = 4$ that is shown on the left-hand side of the middle panel in Figure 3.2. Several quarters later, disagreement reverts back to its pre-crisis level. This suggests that the sustained high level of average individual IU is unlikely the result of model uncertainty alone. Next, we examine the differences among the survey-based IU measures more closely.

3.5.2 Estimates of the Horizon Effect on the Variance of Aggregate Shocks

Table 3.4 contains the estimates of Eqn. (3.10) for a pooled sample including data for all available forecast horizons $h = 4, 8, 20$.

Table 3.4: Estimates of the horizon effect on the variance of aggregate shocks

		Variance of aggregate shocks $\sigma^2_{\lambda,t+h\|t}$
		$h = 4, 8, 20$
$D^{h=4}$	Indicator variable for $h = 4$	1.57
		(0.81)
$D^{h=8}$	Indicator variable for $h = 8$	9.33*
		(0.85)
$D^{h=20}$	Indicator variable for $h = 20$	15.06*
		(1.13)
$t-1$	Time trend	0.37*
		(0.02)
D^{REC}_{t-1}	Recession indicator variable	1.26
		(0.74)
	Adj. R^2	0.96
	No. of observations	164

Notes: This table reports the estimates of Eqn. (3.10). Coefficients are estimated with OLS. Heteroskedasticity-robust standard errors in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks ('*') indicate significance at the 5% critical level. The sample period $t = 1, \dots, 58$ represents time instances between 1999Q1 and 2013Q2.

The results for the horizon-specific indicator variables, $D^{h=4}, D^{h=8}$ and $D^{h=20}$, show that the variance of aggregate shocks increases with h. This underscores the claim of the model of Lahiri and Sheng (2010) that horizon-specific considerations explain a considerable fraction of the difference between IU and disagreement. As can be seen in Figure 3.2, average individual IU increases with h. This might be due to the fact that, five years in advance, the amount of useful information is rather limited. In turn, if disagreement stems from differential interpretation of incoming news, one should also expect low disagreement as long as few signals are available. In the framework of a Bayesian learning model, Lahiri and Sheng (2008) find that distinct interpretations of signals are one of the most important sources of disagreement. In contrast to the horizon effects, the coefficient on D^{REC}_{t-1} is insignificant although it has the positive sign that is predicted in Lahiri and Sheng (2010).

3.5.3 Dynamic Panel Estimates for Individual IU Measures

In this section, we discuss how distinct IU statistics can be explained by macroeconomic conditions and policy variables. Table 3.5 presents the dynamic panel estimates of

Eqn. (3.11).[12] As shown in Eqn. (3.20), the impact of the covariates in $\boldsymbol{m}_{t-1}$ and $\boldsymbol{p}_{t-1}$ on $\sigma^2_{i,t+h|t}$ ($\sigma^2_{\varepsilon,i,t+h|t}$) is equivalent to their effect on $\overline{\sigma^2_{t+h|t}}$ ($s^2_{t+h|t}$). To highlight the distinction between effects of macroeconomic variables on the one hand and policy-related covariates on the other hand, the models are first estimated leaving aside the policy indicators in $\boldsymbol{p}_{t-1}$. These results are shown in the odd-numbered columns.

From Table 3.5 it can be seen that the inflation rate, π_{t-1}, is positively related to $\sigma^2_{i,t+h|t}$. This is in line with the theoretical arguments by Friedman (1977) and Ball (1992). However, the significance of the relation is statistically weak. The association between the inflation rate and IU might be less pronounced than in earlier studies since the relationship hypothesized by Friedman-Ball refers mainly to high-inflation regimes, where IU rises due to uncertainty about the timing of disinflation policies adopted by a central bank. Hence, the low-inflation environment that characterizes most of the sample period for the Euro area might be one reason for the relatively small effect. This is in line with the findings of Baillie et al. (1996). The insignificant coefficients in the even-numbered columns may also be explained by the fact that π_{t-1} is positively correlated with both TD^+_{t-1} and $|TD^-_{t-1}|$ due to the inclusion of the inflation rate as a covariate in Eqn. (3.18). In addition, π_{t-1} is positively correlated with ΔMPC_{t-1}, reflecting the fact that higher values of the Monetary Policy Communicator imply that the central bank communicates inflation risks. In line with the arguments of Friedman (1977) that higher economic growth reduces IU, the results indicate a significant negative relationship between $\sigma^2_{i,t+h|t}$ and Δx_{t-1}. Hence, IU rises during economic downturns and vice versa. This is in line with recent findings, e.g., of Bloom (2009). Stock market volatility is insignificant for all horizons, though the estimated coefficients have the expected positive sign for $h = 4$ and 8.

Turning to the relationship between IU and indicators of monetary policy in the even-numbered columns, we find that during periods of expansionary monetary policy such as the years following the beginning of the crisis, (average) individual IU also tends to be high. This is expressed by the positive relationship between negative deviations from the Taylor rule, $|TD^-_{t-1}|$, and $\sigma^2_{i,t+h|t}$ for all h and highlights the important role of monetary policy for the emergence of IU. High uncertainty during periods of expansionary monetary policy might emerge in cases when the central bank aims to increase inflation, yet survey participants are uncertain about the success of such attempts. As can be seen in Figure 3.2, the $\overline{\sigma^2_{t+h|t}}$ statistics increase from an initial level of around 0.2 before the beginning of the crisis to approximately 0.4 afterwards. The interest rate policy of the ECB during the crisis is characterized by sustained periods of negative deviations

[12] The estimates based on the alternative definitions of D^{REC}_{t-1}, ΔAS_{t-1} or TD_{t-1} described in Footnotes 5, 7, and 8 are presented in Tables 3.7-3.9 in the Appendix.

Table 3.5: Dynamic panel estimates of the relationship between individual IU measures and their covariates

		Individual IU $\sigma^2_{i,t+h\|t}$						Variance of individual shocks $\sigma^2_{\varepsilon,i,t+h\|t}$					
		$h=4$		$h=8$		$h=20$		$h=4$		$h=8$		$h=20$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sigma^2_{i,t+h-1\|t-1}$	Individual IU	38.29*	36.52*	40.88*	38.04*	54.67*	47.33*						
		(7.47)	(6.00)	(8.59)	(7.65)	(9.57)	(8.71)						
$\sigma^2_{\varepsilon,i,t+h-1\|t-1}$	Variance of individual shocks							41.09*	40.10*	32.29*	39.82*	39.00*	40.91*
								(7.87)	(6.52)	(10.31)	(8.25)	(10.11)	(8.49)
π_{t-1}	Inflation rate	1.86	0.58	2.54*	0.54	1.84	0.67	3.90*	0.28	2.31	−0.96	1.84	−0.09
		(1.01)	(0.61)	(1.05)	(0.76)	(1.24)	(1.12)	(1.19)	(0.69)	(1.19)	(0.95)	(1.66)	(1.08)
Δx_{t-1}	Real GDP growth rate	−1.42*	−1.62*	−1.54*	−1.65*	−1.43*	−1.88*	−1.84*	−1.11*	−2.48*	−1.01	−2.17*	−1.22*
		(0.69)	(0.47)	(0.60)	(0.51)	(0.71)	(0.62)	(0.81)	(0.53)	(0.68)	(0.53)	(0.85)	(0.59)
$RV_{t-1}(E)$	Stock price volatility	5.06	7.00	1.95	5.64	−8.12	−10.55	32.05*	39.51*	4.22	25.26*	8.83	7.35
		(7.01)	(8.14)	(5.66)	(8.49)	(8.55)	(10.78)	(8.86)	(9.31)	(6.81)	(9.98)	(9.72)	(11.26)
ΔEPU_{t-1}	Economic Policy Uncertainty		−0.16		−0.18		−0.35		−0.54		−3.44		−1.00
			(1.63)		(1.90)		(2.41)		(1.95)		(2.35)		(2.29)
ΔAS_{t-1}	ECB assets		−4.18		−5.54		0.99		10.21*		−2.46		−3.40
			(4.05)		(4.54)		(4.68)		(4.60)		(5.31)		(5.05)
TD^{+}_{t-1}	Positive Taylor deviations		3.01*		1.52		−2.76		5.61*		4.67*		−2.37
			(1.26)		(1.62)		(2.13)		(1.39)		(2.05)		(2.19)
$\|TD^{-}_{t-1}\|$	Negative Taylor deviations		3.96*		4.47*		3.72*		0.71		0.92		1.07
			(0.95)		(1.31)		(1.12)		(1.01)		(1.42)		(1.13)
MPC_{t-1}	Monetary Policy Communicator		2.25		0.90		2.24		3.00		0.37		2.63
			(2.48)		(2.48)		(3.85)		(3.15)		(2.99)		(3.86)
ΔMM_{t-1}	Money multiplier		−0.27		−0.92		0.15		2.69*		0.99		−1.26
			(1.00)		(1.05)		(1.19)		(1.15)		(1.26)		(1.29)
	Constant	6.81*	9.26*	8.44*	13.40*	9.97*	19.30*	−1.35	−0.68	7.60*	3.61	5.22	6.23
		(2.05)	(2.28)	(2.93)	(3.40)	(4.13)	(4.47)	(2.26)	(2.36)	(3.28)	(3.32)	(5.16)	(4.92)
$t-1$	Time trend	0.12*	0.05	0.14*	0.01	0.13	−0.05	−0.02	0.04	−0.09	−0.00	−0.09	−0.08
		(0.05)	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.09)	(0.10)	(0.09)
D^{REC}_{t-1}	Recession indicator variable	−1.51	−0.32	−1.71	−0.01	−0.92	−1.13	−3.79	−0.15	−4.81*	−0.34	−4.11	−1.60
		(1.81)	(1.06)	(1.76)	(1.23)	(1.95)	(1.44)	(2.16)	(1.15)	(2.07)	(1.45)	(2.60)	(1.56)
	No. of observations	2,206	2,157	1,928	1,880	1,533	1,533	2,206	2,157	1,928	1,880	1,533	1,533
	No. of forecasters	88	87	87	86	75	75	88	87	87	86	75	75
	No. of instruments	32	74	32	74	32	68	32	74	32	74	32	68
	Hansen p-value	0.09	0.42	0.17	0.79	0.12	0.37	0.00	0.14	0.13	0.33	0.04	0.39

Notes: This table reports the estimates of Eqn. (3.11). Coefficients are estimated with the two-step system GMM estimator by Blundell and Bond (1998). Within-forecaster heteroskedasticity- and autocorrelation-robust standard errors using the adjustment proposed by Windmeijer (2005) are reported in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks (‘*’) indicate significance at the 5% critical level. Cross-sectional units $i = 1, \ldots, 98$ denote individual forecasters and sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

from the optimal Taylor rule of up to three percentage points, which can be seen in Figure 3.3. The estimated coefficients on $|TD^{-}_{t-1}|$ in columns (2), (4) and (6) suggest that a downward deviation of the interest rate from the optimal Taylor rule by one percentage point increases (average) individual IU by approximately 0.04 units. This is a relatively large effect and suggests that expansionary monetary policy accounts for more than half of the increase in average individual IU since the beginning of the crisis for all horizons. In contrast, positive deviations from the Taylor rule, TD^{+}_{t-1}, are only related to short-term IU, $\sigma^2_{i,t+4|t}$, although the estimated effect is also relatively large.

In columns (7) to (12) we present the estimates when the variance of individual shocks, $\sigma^2_{\varepsilon,i,t+h|t}$, is used as the dependent variable. The relationship between macroeconomic conditions and $\sigma^2_{\varepsilon,i,t+h|t}$ is similar to the evidence for $\sigma^2_{i,t+h|t}$. However, in line with Engle and Rangel (2008) and Conrad and Loch (2015), realized stock market volatility, $RV_{t-1}(E)$ is positively related to the variance of individual shocks for $h = 4$ and 8. A possible explanation for this finding might be the heterogeneity in the panel of forecasters, which includes both financial institutions and research institutes. The former may adjust their predictions in a distinct way during periods of increased volatility in stock markets. In such cases, it is possible that all forecasters retain their individual uncertainty but the disagreement increases. However, due to the anonymity of the participants in the SPF panel, it is difficult to test this hypothesis empirically. With respect to the importance of monetary policy, we find that $\sigma^2_{\varepsilon,i,t+h|t}$ is positively related to TD^{+}_{t-1} for $h = 4$ and 8. This implies that disagreement rises during periods when the interest rate is higher, and thus more restrictive, than the one prescribed by a Taylor rule. Periods of contractionary monetary policy are expected to lead to a decline in the inflation rate, which should reduce disagreement based on the positive relationship between π_{t-1} and $\sigma^2_{\varepsilon,i,t+h|t}$. However, the positive impact of TD^{+}_{t-1} suggests that the decision of the central bank to deviate from a rules-based monetary policy increases disagreement in a way that outweighs the potential reduction that results from the policy-induced decrease of the inflation rate. Since disagreement is a component of average individual IU and since contractionary monetary policy is not related to individual IU, the overall effect of such deviations on IU appears to be small, despite its relationship with disagreement. Rather, it is periods of expansionary interest rate policy that are related to increases in IU. We also find that $\sigma^2_{\varepsilon,i,t+4|t}$ is positively related to ΔAS_{t-1} and ΔMM_{t-1}, which implies that forecasters disagree about the implications of large changes in the ECB's balance sheet and monetary aggregates for future inflation in the short, but not in the long run. To summarize, individual IU and the variance of individual shocks are related to distinct indicators of monetary policy. In particular, forecasters disagree strongly about the effects of contractionary monetary policy, but remain individually confident in such situations.[13] In contrast, forecasters are uncertain about the impact of expansionary monetary policy. This asymmetry affects assessments of whether unpredictable changes in monetary policy are associated with increases in the level of IU or not. Given the results for $\sigma^2_{\varepsilon,i,t+h|t}$ from Table 3.5, it appears that positive deviations from a rules-based interest rate policy are associated with higher IU. However, it is negative deviations that are most important for explaining the sustained increase in average individual IU across all forecast horizons since roughly the beginning of the crisis.

To examine the validity of the overidentifying restrictions, Table 3.5 also reports p-values for the Hansen test. We cannot reject the null hypothesis in all model specifications for $\sigma^2_{i,t+h|t}$ and almost all cases for the variance of individual shocks, $\sigma^2_{\varepsilon,i,t+h|t}$. This supports the robustness of the employed specification. The number of instrument variables is moderate and smaller than the cross-sectional dimension N.

3.5.4 Estimates of the Covariates of the Variance of Aggregate Shocks

The results in the previous section reveal that different measures of IU are related to distinct indicators of monetary policy. In particular, average individual IU rises during periods of expansionary monetary policy, while disagreement does not. In order to examine this asymmetric effect in more detail, we relate $\sigma^2_{\lambda,t+h|t}$ to the predictor variables in $\boldsymbol{m}_{t-1}$ and $\boldsymbol{p}_{t-1}$. If $\sigma^2_{\lambda,t+h|t}$ can be explained by these factors, the results from Table 3.5 have to differ between $\overline{\sigma^2_{t+h|t}}$ and $s^2_{t+h|t}$. In particular, it would suggest that the asymmetric effect of monetary policy on different measures of IU is not coincidental, but rather a characteristic of the variance of the aggregate shocks that are unpredictable ex-ante. The estimates of Eqn. (3.21) are reported in Table 3.6.

We find that $\sigma^2_{\lambda,t+h|t}$ is more closely related to the covariates, i.e., more predictable, for large h. In particular, the perceived variance of aggregate shocks is negatively related to most indicators of monetary policy. Thus, it appears that distinct monetary policy instruments available to the ECB successfully reduce the part of IU that is common across forecasters. However, $\sigma^2_{\lambda,t+h|t}$ is *positively* related to the deviation variable $|TD^-_{t-1}|$. This means that periods of expansionary monetary policy increase the perceived variance of the forthcoming aggregate inflationary shocks. A possible explanation for this is that forecasters believe that the ECB's monetary policy since the beginning of the crisis has a

[13] Using alternative measures of TD as discussed in Footnote 8, we have found a significant negative relationship between $\sigma^2_{i,t+20|t}$ and TD^+_{t-1} (see Table 3.9). This implies that contractionary monetary policy today *reduces* individual uncertainty regarding inflation in five years.

Table 3.6: Estimates of the covariates of the variance of aggregate shocks

		Variance of agg. shocks $\sigma^2_{\lambda,t+h\|t}$		
		$h=4$	$h=8$	$h=20$
		(1)	(2)	(3)
π_{t-1}	Inflation rate	1.99	3.75*	3.27*
		(1.26)	(1.13)	(0.93)
Δx_{t-1}	Real GDP growth rate	−0.65	−1.34*	−1.45*
		(0.40)	(0.59)	(0.47)
$RV_{t-1}(E)$	Stock price volatility	−49.66	−24.41	−34.91
		(29.72)	(19.98)	(22.78)
ΔEPU_{t-1}	Economic Policy Uncertainty	5.62	4.47	4.53
		(4.13)	(3.75)	(4.01)
ΔAS_{t-1}	ECB assets	−25.48*	−9.96	−6.42
		(10.48)	(8.55)	(12.03)
TD^+_{t-1}	Positive Taylor deviations	−3.81	−4.12	−6.98*
		(2.45)	(2.42)	(2.64)
$\|TD^-_{t-1}\|$	Negative Taylor deviations	4.87*	5.71*	4.13*
		(1.59)	(0.97)	(1.07)
MPC_{t-1}	Monetary Policy Communicator	−4.81	0.07	−7.07*
		(3.60)	(2.40)	(2.79)
ΔMM_{t-1}	Money multiplier	−4.86*	−3.62*	−1.67
		(1.59)	(1.26)	(2.41)
	Constant	12.23*	12.87*	20.64*
		(3.55)	(2.80)	(3.75)
$t-1$	Time trend	0.06	0.07	0.14*
		(0.09)	(0.06)	(0.07)
D^{REC}_{t-1}	Recession indicator variable	−0.90	0.17	−0.95
		(1.62)	(1.41)	(1.64)
	Adj. R^2	0.66	0.81	0.73
	No. of observations	56	56	50

Notes: This table reports the estimates of Eqn. (3.21). Coefficients are estimated with OLS. Newey-West standard errors accounting for fourth-order autocorrelation in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks (‘*’) indicate significance at the 5% critical level. The sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

destabilizing effect on the economy of the Euro area. For example, forecasters might be uncertain about the success of the ECB's attempts to increase inflation towards the target of around 2%. This finding helps to explain the previous evidence regarding the covariates of IU from Table 3.5. For example, positive deviations from the Taylor rule increase disagreement but systematically reduce the perceived variance of aggregate shocks. If these effects offset each other, average individual IU may not appear to be significantly related to TD^{+}_{t-1}. In contrast, negative deviations increase the perceived variance of aggregate shocks but seem to have little impact on the variance of individual shocks and, therefore, disagreement. Hence, the overall effect on average individual IU is positive. In other words, $|TD^{-}_{t-1}|$ explains the difference between $\overline{\sigma^2_{t+h|t}}$ and $s^2_{t+h|t}$ and, thereby, increases average individual IU. Hence, policy evaluations based on disagreement as a proxy for IU are potentially incomplete and potentially misleading.

3.6 Conclusion

We analyze average individual IU and disagreement as two distinct measures of IU. The estimation sample covers the periods before and after the beginning of the financial crisis in the Eurozone. While disagreement reverts to its pre-crisis level after a short period of time, average individual IU continues to rise. In line with the decomposition proposed by Lahiri and Sheng (2010), which shows that average individual IU can be regarded as the sum of disagreement and the variance of aggregate shocks, the empirical evidence shows that the difference between average individual IU and disagreement increases with the forecast horizon. Based on the empirical confirmation of this decomposition, we relate different measures of IU to a number of macroeconomic and policy variables. Based on a cross-section of forecasters, we find that the relationship between IU and macroeconomic conditions is in line with typical findings in the literature. However, the employed measures of IU are related to fundamentally distinct indicators of monetary policy. Most strikingly, average individual IU is primarily associated with expansionary monetary policy, whereas disagreement is rather related to contractionary monetary policy. Moreover, we find that the difference between these measures systematically increases when monetary policy is expansive. Thus, disagreement is a misleading measure of the risks associated with the European Central Bank's management of the financial and sovereign debt crisis.

3.7 Appendix

Below we report the estimates of Eqn. (3.11) for alternative choices of D_{t-1}^{REC}, ΔAS_{t-1} and TD_{t-1} as described in Footnotes 5, 7, and 8, respectively.

Table 3.7: Dynamic panel estimates of the relationship between individual IU measures and their covariates (alternative D_{t-1}^{REC})

		Individual IU $\sigma^2_{i,t+h\|t}$						Variance of individual shocks $\sigma^2_{\varepsilon,i,t+h\|t}$					
		$h=4$		$h=8$		$h=20$		$h=4$		$h=8$		$h=20$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sigma^2_{i,t+h-1\|t-1}$	Individual IU	39.20*	36.35*	40.54*	37.50*	47.01*	47.99*						
		(6.86)	(5.89)	(8.62)	(8.73)	(8.96)	(8.68)						
$\sigma^2_{\varepsilon,i,t+h-1\|t-1}$	Variance of individual shocks							41.48*	39.95*	32.10*	39.65*	37.24*	41.79*
								(7.72)	(6.62)	(9.76)	(8.36)	(9.79)	(8.49)
π_{t-1}	Inflation rate	1.66*	0.76	2.20*	0.70	1.03	−0.13	3.13*	0.06	1.36	−0.76	−0.05	−0.72
		(0.65)	(0.76)	(0.83)	(0.82)	(0.84)	(1.06)	(0.83)	(0.87)	(1.14)	(1.02)	(1.26)	(1.13)
Δx_{t-1}	Real GDP growth rate	−0.72	−1.04*	−0.68	−1.41*	−3.13*	−1.97*	−0.92	−1.22*	−0.76	−0.59	−2.11*	−0.38
		(0.46)	(0.41)	(0.56)	(0.45)	(0.75)	(0.55)	(0.57)	(0.49)	(0.55)	(0.47)	(0.96)	(0.58)
$RV_{t-1}(E)$	Stock price volatility	3.56	6.03	0.29	6.80	−13.71	−11.16	27.38*	40.30*	−2.07	24.41*	−1.88	3.74
		(6.54)	(8.06)	(5.79)	(8.90)	(8.04)	(10.22)	(8.94)	(9.71)	(7.32)	(10.04)	(10.22)	(11.16)
ΔEPU_{t-1}	Economic Policy Uncertainty		0.01		−0.53		−0.26		−0.93		−2.96		−0.30
			(1.56)		(1.92)		(2.29)		(2.00)		(2.36)		(2.33)
ΔAS_{t-1}	ECB assets		−5.75		−4.86		0.48		9.53*		−4.15		−6.91
			(3.96)		(4.46)		(4.91)		(4.26)		(5.24)		(5.27)
TD^+_{t-1}	Positive Taylor deviations		2.19		1.17		−2.62		5.37*		3.82		−4.00
			(1.28)		(1.75)		(2.21)		(1.51)		(2.06)		(2.38)
$\|TD^-_{t-1}\|$	Negative Taylor deviations		3.94*		4.20*		3.06*		0.28		0.88		1.02
			(1.06)		(1.36)		(1.17)		(1.07)		(1.45)		(1.23)
MPC_{t-1}	Monetary Policy Communicator		1.40		0.69		4.07		3.64		−0.58		3.08
			(2.47)		(2.47)		(3.96)		(3.18)		(3.06)		(3.82)
ΔMM_{t-1}	Money multiplier		−0.67		−0.82		0.15		2.75*		0.59		−2.31
			(1.00)		(1.05)		(1.30)		(1.18)		(1.24)		(1.42)
	Constant	4.95*	7.76*	6.57*	12.16*	14.02*	19.39*	−3.43	−0.49	3.57	2.57	4.63	5.72
		(2.12)	(2.14)	(3.22)	(3.29)	(4.02)	(4.24)	(2.54)	(2.31)	(3.41)	(3.20)	(5.10)	(4.50)
$t-1$	Time trend	0.12*	0.05	0.15*	0.02	0.26*	0.01	−0.00	0.07	−0.06	−0.00	0.04	−0.07
		(0.05)	(0.06)	(0.06)	(0.07)	(0.08)	(0.09)	(0.05)	(0.06)	(0.06)	(0.09)	(0.11)	(0.11)
D^{REC}_{t-1}	Recession indicator variable	3.07	2.75	3.55	0.63	−11.29*	−3.45	1.09	−1.90	4.47	1.43	−6.27	1.08
		(2.61)	(1.82)	(3.05)	(1.87)	(3.54)	(2.81)	(3.06)	(2.20)	(3.49)	(2.17)	(4.78)	(3.09)
	No. of observations	2,206	2,157	1,928	1,880	1,533	1,533	2,206	2,157	1,928	1,880	1,533	1,533
	No. of forecasters	88	87	87	86	75	75	88	87	87	86	75	75
	No. of instruments	32	74	32	74	32	68	32	74	32	74	32	68
	Hansen p-value	0.09	0.41	0.24	0.78	0.30	0.41	0.00	0.17	0.10	0.33	0.03	0.40

Notes: This table reports the estimates of Eqn. (3.11) when a recession is defined as two (or more) consecutive periods characterized by a negative real GDP growth rate. Coefficients are estimated with the two-step system GMM estimator by Blundell and Bond (1998). Within-forecaster heteroskedasticity- and autocorrelation-robust standard errors using the adjustment proposed by Windmeijer (2005) are reported in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks (‘*’) indicate significance at the 5% critical level. Cross-sectional units $i = 1, \ldots, 98$ denote individual forecasters and sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

Table 3.8: Dynamic panel estimates of the relationship between individual IU measures and their covariates (alternative ΔAS_{t-1})

		Individual IU $\sigma^2_{i,t+h\|t}$						Variance of individual shocks $\sigma^2_{\varepsilon,i,t+h\|t}$					
		$h=4$		$h=8$		$h=20$		$h=4$		$h=8$		$h=20$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sigma^2_{i,t+h-1\|t-1}$	Individual IU	38.29*	36.80*	40.88*	38.21*	54.67*	46.93*						
		(7.47)	(6.07)	(8.59)	(7.64)	(9.57)	(8.72)						
$\sigma^2_{\varepsilon,i,t+h-1\|t-1}$	Variance of individual shocks							41.09*	39.53*	32.29*	39.35*	39.00*	40.78*
								(7.87)	(6.50)	(10.31)	(8.32)	(10.11)	(8.46)
π_{t-1}	Inflation rate	1.86	0.37	2.54*	0.48	1.84	0.76	3.90*		0.04 2.31	−1.12	1.84	−0.03
		(1.01)	(0.60)	(1.05)	(0.77)	(1.24)	(1.12)	(1.19)	(0.66)	(1.19)	(0.94)	(1.66)	(1.08)
Δx_{t-1}	Real GDP growth rate	−1.42*	−1.55*	−1.54*	−1.65*	−1.43*	−1.90*	−1.84*	−1.00	−2.48*	−1.06*	−2.17*	−1.30*
		(0.69)	(0.48)	(0.60)	(0.50)	(0.71)	(0.62)	(0.81)	(0.53)	(0.68)	(0.53)	(0.85)	(0.59)
$RV_{t-1}(E)$	Stock price volatility	5.06	7.27	1.95	5.31	−8.12	−11.32	32.05*	38.56*	4.22	24.73*	8.83	5.96
		(7.01)	(8.18)	(5.66)	(8.55)	(8.55)	(10.63)	(8.86)	(9.26)	(6.81)	(10.08)	(9.72)	(11.13)
ΔEPU_{t-1}	Economic Policy Uncertainty		−0.29		−0.20		−0.10		−0.73		−3.76		−0.97
			(1.66)		(1.97)		(2.40)		(1.95)		(2.40)		(2.29)
ΔAS_{t-1}	ECB assets		−8.22		−11.36		0.52		31.62*		5.86		−1.53
			(9.76)		(10.76)		(11.27)		(10.74)		(12.57)		(11.98)
TD^+_{t-1}	Positive Taylor deviations		3.28*		1.67		−2.55		5.72*		4.73*		−2.19
			(1.24)		(1.59)		(2.10)		(1.39)		(2.02)		(2.17)
$\|TD^-_{t-1}\|$	Negative Taylor deviations		3.88*		4.41*		3.83*		0.59		0.71		1.03
			(0.97)		(1.30)		(1.12)		(1.02)		(1.45)		(1.13)
MPC_{t-1}	Monetary Policy Communicator		2.59		0.98		2.22		2.96		0.16		2.45
			(2.51)		(2.48)		(3.90)		(3.09)		(2.98)		(3.89)
ΔMM_{t-1}	Money multiplier		−0.19		−0.84		−0.01		3.28*		1.76		−0.78
			(1.06)		(1.10)		(1.24)		(1.19)		(1.33)		(1.33)
	Constant	6.81*	9.07*	8.44*	13.35*	9.97*	19.37*	−1.35	−0.59	7.60*	4.20	5.22	6.58
		(2.05)	(2.29)	(2.93)	(3.37)	(4.13)	(4.45)	(2.26)	(2.34)	(3.28)	(3.31)	(5.16)	(4.88)
$t-1$	Time trend	0.12*	0.06	0.14*	0.01	0.13	−0.06	−0.02	0.05	−0.09	0.00	−0.09	−0.08
		(0.05)	(0.06)	(0.06)	(0.07)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.09)	(0.10)	(0.09)
D^{REC}_{t-1}	Recession indicator variable	−1.51	−0.05	−1.71	0.09	−0.92	−1.16	−3.79	0.02	−4.81*	−0.57	−4.11	−1.78
		(1.81)	(1.04)	(1.76)	(1.17)	(1.95)	(1.43)	(2.16)	(1.14)	(2.07)	(1.37)	(2.60)	(1.55)
	No. of observations	2,206	2,157	1,928	1,880	1,533	1,533	2,206	2,157	1,928	1,880	1,533	1,533
	No. of forecasters	88	87	87	86	75	75	88	87	87	86	75	75
	No. of instruments	32	74	32	74	32	68	32	74	32	74	32	68
	Hansen p-value	0.09	0.38	0.17	0.79	0.12	0.38	0.00	0.14	0.13	0.33	0.04	0.38

Notes: This table reports the estimates of Eqn. (3.11) when we replace ΔAS_{t-1} with the changes of the ECB assets-to-nominal-GDP ratio. Coefficients are estimated with the two-step system GMM estimator by Blundell and Bond (1998). Within-forecaster heteroskedasticity- and autocorrelation-robust standard errors using the adjustment proposed by Windmeijer (2005) are reported in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks ('*') indicate significance at the 5% critical level. Cross-sectional units $i = 1, \ldots, 98$ denote individual forecasters and sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

Table 3.9: Dynamic panel estimates of the relationship between individual IU measures and their covariates (alternative TD_{t-1})

		Individual IU $\sigma^2_{i,t+h\|t}$						Variance of individual shocks $\sigma^2_{\varepsilon,i,t+h\|t}$					
		$h=4$		$h=8$		$h=20$		$h=4$		$h=8$		$h=20$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sigma^2_{i,t+h-1\|t-1}$	Individual IU	38.29*	38.45*	40.88*	38.11*	54.67*	50.34*						
		(7.47)	(6.21)	(8.59)	(7.59)	(9.57)	(8.43)						
$\sigma^2_{\varepsilon,i,t+h-1\|t-1}$	Variance of individual shocks							41.09*	38.22*	32.29*	39.92*	39.00*	42.52*
								(7.87)	(6.41)	(10.31)	(8.60)	(10.11)	(8.35)
π_{t-1}	Inflation rate	1.86	0.60	2.54*	1.01	1.84	−0.49	3.90*	0.97	2.31	0.18	1.84	−0.55
		(1.01)	(0.55)	(1.05)	(0.57)	(1.24)	(1.03)	(1.19)	(0.62)	(1.19)	(0.74)	(1.66)	(1.04)
Δx_{t-1}	Real GDP growth rate	−1.42*	−1.46*	−1.54*	−1.58*	−1.43*	−1.18*	−1.84*	−1.46*	−2.48*	−1.42*	−2.17*	−0.93
		(0.69)	(0.42)	(0.60)	(0.39)	(0.71)	(0.48)	(0.81)	(0.47)	(0.68)	(0.47)	(0.85)	(0.49)
$RV_{t-1}(E)$	Stock price volatility	5.06	4.50	1.95	2.85	−8.12	−18.42	32.05*	39.52*	4.22	27.20*	8.83	5.78
		(7.01)	(8.46)	(5.66)	(9.55)	(8.55)	(11.25)	(8.86)	(9.23)	(6.81)	(11.54)	(9.72)	(12.12)
ΔEPU_{t-1}	Economic Policy Uncertainty		−0.27		−0.69		0.37		−0.84		−3.70		−1.53
			(1.60)		(1.90)		(2.41)		(1.88)		(2.51)		(2.36)
ΔAS_{t-1}	ECB assets		−5.89		−4.41		1.43		7.49		−2.57		−3.91
			(3.99)		(4.30)		(4.21)		(4.47)		(5.13)		(4.58)
TD^+_{t-1}	Positive Taylor deviations		3.73*		1.37		−5.02*		7.62*		4.48*		−3.34
			(1.27)		(1.49)		(1.93)		(1.34)		(1.73)		(2.14)
$\|TD^-_{t-1}\|$	Negative Taylor deviations		4.58*		4.55*		3.11*		1.83		0.74		0.16
			(1.10)		(1.37)		(1.15)		(1.07)		(1.62)		(1.12)
MPC_{t-1}	Monetary Policy Communicator		1.03		−0.86		3.35		3.73		−0.10		3.16
			(2.41)		(2.37)		(3.78)		(2.96)		(3.11)		(3.84)
ΔMM_{t-1}	Money multiplier		−0.38		−0.70		−0.31		2.54*		1.05		−1.96
			(0.99)		(0.98)		(1.14)		(1.11)		(1.33)		(1.21)
	Constant	6.81*	9.21*	8.44*	12.43*	9.97*	19.55*	−1.35	−0.78	7.60*	3.13	5.22	5.61
		(2.05)	(2.23)	(2.93)	(3.30)	(4.13)	(4.43)	(2.26)	(2.22)	(3.28)	(3.72)	(5.16)	(4.92)
$t-1$	Time trend	0.12*	0.04	0.14*	0.01	0.13	−0.02	−0.02	−0.01	−0.09	−0.02	−0.09	−0.02
		(0.05)	(0.05)	(0.06)	(0.07)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.09)	(0.10)	(0.09)
D^{REC}_{t-1}	Recession indicator variable	−1.51	−0.70	−1.71	−0.96	−0.92	−0.17	−3.79	−0.83	−4.81*	−1.78	−4.11	−1.08
		(1.81)	(1.00)	(1.76)	(1.10)	(1.95)	(1.22)	(2.16)	(1.08)	(2.07)	(1.47)	(2.60)	(1.37)
	No. of observations	2,206	2,157	1,928	1,880	1,533	1,533	2,206	2,157	1,928	1,880	1,533	1,533
	No. of forecasters	88	87	87	86	75	75	88	87	87	86	75	75
	No. of instruments	32	74	32	74	32	68	32	74	32	74	32	68
	Hansen p-value	0.09	0.39	0.17	0.83	0.12	0.38	0.00	0.16	0.13	0.24	0.04	0.39

Notes: This table reports the estimates of Eqn. (3.11) when we assume a two-year horizon in Eqn. (3.18) and use the estimates to construct TD^+_{t-1} and $|TD^-_{t-1}|$. Coefficients are estimated with the two-step system GMM estimator by Blundell and Bond (1998). Within-forecaster heteroskedasticity- and autocorrelation-robust standard errors using the adjustment proposed by Windmeijer (2005) are reported in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks ('*') indicate significance at the 5% critical level. Cross-sectional units $i = 1, \ldots, 98$ denote individual forecasters and sample period $t = 1, \ldots, 58$ represents time instances between 1999Q1 and 2013Q2.

Chapter 4

Five Dimensions of the Uncertainty-Disagreement Linkage

4.1 Introduction

The negative influence of uncertainty shocks on real economic activity is well documented in the literature (Jurado et al., 2015; Meinen and Röhe, 2017; Jo and Sekkel, 2018). Hence, macroeconomic forecasts should be complemented by some quantification of the underlying uncertainty. Two of the most commonly used survey-based indicators are the cross-sectional dispersion of the point predictions (henceforth 'disagreement') and the average across the second moments of the histogram forecasts ('uncertainty'). We analyze the relationship between disagreement and uncertainty in the Euro area using data from the Survey of Professional Forecasters (SPF) conducted by the European Central Bank (ECB) for the period 1999Q1–2018Q2.

The widespread use of disagreement as an indicator of uncertainty in earlier studies is primarily ascribed to the unavailability of density forecasts. Dovern (2015) finds that forecaster disagreement is positively related to realized stock market volatility and measures of uncertainty based on newspaper articles (cf. Baker et al., 2016). In addition to the point predictions, several surveys of macroeconomic expectations including the SPF elicit probability distributions in the form of histograms from their participants. These forecasts can be used to construct direct measures of ex-ante uncertainty and to empirically examine the relationship between uncertainty and disagreement. The evidence regarding the strength of this link is mixed. Early research by Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003) suggests that disagreement is a good proxy for uncertainty. More recent studies by Lahiri and Liu (2006), D'Amico and Orphanides (2008), Rich and

This chapter is based on a single-authored working paper of the same name (cf. Glas, 2018).

Tracy (2010) and Glas and Hartmann (2016) find contrary evidence. However, each of the articles listed above relies on a number of assumptions that may have a considerable impact on the strength of the link between uncertainty and disagreement (Grimme et al., 2014; Abel et al., 2016). We contribute to the literature by conducting an empirical analysis that accounts for the following aspects:

Dispersion statistic: We consider distinct statistics that differ in terms of their robustness to extreme observations. Most studies rely on variance-based indicators, which may be strongly affected by outliers (Lahiri and Sheng, 2010; Glas and Hartmann, 2016). This is practically relevant because the cross-section of survey participants is relatively small in many cases. Therefore, we consider measures based on the interquartile range in addition to the variance-based statistics.

Disagreement measure: Forecaster disagreement can be calculated using either the point forecasts or the means of the reported histograms. However, it has been frequently observed that the two series tend to deviate in practice (Clements, 2010, 2012). The choice between the two may thus have a considerable impact on the observed cross-sectional dispersion of the predictions. Nonetheless, most studies focus on either the point predictions (Zarnowitz and Lambros, 1987; Giordani and Söderlind, 2003; Boero et al., 2015) or the histogram means (Glas and Hartmann, 2016), thereby ignoring the other. In order to assess how strongly the link between uncertainty and disagreement is affected by this choice, we use both.

Distributional assumptions: We account for various commonly used assumptions regarding the distribution of the probability mass within the distinct outcome intervals of the survey questionnaire. This includes the use of parametric distributions. Many studies rely on a single assumption. As discussed by Giordani and Söderlind (2003), this choice has an impact on the observed level of uncertainty.

Outcome variables: Most studies that analyze the SPF data focus on either the inflation rate (D'Amico and Orphanides, 2008; Glas and Hartmann, 2016), output growth (Clements, 2011), or both (Zarnowitz and Lambros, 1987; Giordani and Söderlind, 2003; Boero et al., 2015). However, the strength of the uncertainty-disagreement linkage may differ for more persistent outcome variables such as the unemployment rate. Hence, differences across outcome variables are considered.

Forecast horizon: Lahiri and Sheng (2010) use a theoretical forecast error decomposition to show that the difference between disagreement and uncertainty increases with the forecast horizon due to an accumulation of aggregate future shocks. Similarly, Boero et al. (2015) show that the relationship between the measures depends on the forecast horizon. Therefore, we evaluate whether uncertainty and disagreement are more closely related at specific forecast horizons.

We find that disagreement is generally a poor proxy for uncertainty. However, we show that the strength of the relationship is related to the previously discussed aspects which are frequently neglected or insufficiently addressed in many papers. In particular, we document that the strength of the association depends on the employed dispersion statistic, the usage of either point forecasts or histogram means to calculate disagreement, the considered outcome variable and the forecast horizon. In contrast, distributional assumptions do not appear to be very influential. Moreover, the relationship is weaker during economically turbulent periods when indicators of uncertainty are needed most. Accounting for the entry and exit of forecasters to and from the SPF panel has little impact on the results. We also document that survey-based uncertainty is associated with overall policy uncertainty, whereas forecaster disagreement is more closely related to the fluctuations on financial markets. The same factors that are responsible for the decoupling of policy and financial uncertainty may thus also explain the divergence between survey-based measures of uncertainty and disagreement.

The most comprehensive previous analysis in terms of addressing the aspects listed above is provided by Abel et al. (2016), who account for the impact of the employed dispersion statistic, the considered outcome variable and the forecast horizon. However, the well-documented difference between the point forecasts and the histogram means as well as the impact of distributional assumptions are disregarded. Moreover, the sample of Abel et al. (2016) does not cover recent politically and economically relevant events such as the decision of the United Kingdom (UK) to leave the European Union (EU) in June 2016. This study extends the analysis of Abel et al. (2016) and accounts for the missing aspects listed above. To our knowledge this is the most comprehensive evaluation of the link between uncertainty and disagreement based on the SPF data.

The rest of this chapter is structured as follows: In Section 4.2, we discuss the survey-based measures of disagreement and uncertainty. The forecast data from the SPF are described in Section 4.3. Section 4.4 presents the estimates of the relationship between disagreement and uncertainty. In Section 4.5 we provide a comparison to other popular measures of uncertainty. Section 4.6 summarizes the findings and concludes.

4.2 Measuring Disagreement and Uncertainty

In this section, we describe the employed measures of disagreement and uncertainty based on survey data. Moreover, we discuss the empirical models that are used to analyze the strength of their relationship.

4.2.1 Forecaster Disagreement

Disagreement among professional forecasters has been widely used as a proxy for uncertainty in the empirical literature due to the fact that only point forecasts are required in the computation. Consider a cross-section of N forecasters, who report h-step-ahead predictions, $\mu_{i,t+h|t}$ for $i = 1, \ldots, N$, in consecutive quarters $t = 1, \ldots, T$ for some future outcome, x_{t+h}. A commonly used measure of disagreement is the cross-sectional standard deviation of the point predictions,

$$s_{t+h|t} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\mu_{i,t+h|t} - \overline{\mu}_{t+h|t})^2}, \tag{4.1}$$

where $\overline{\mu}_{t+h|t} = (1/N)\sum_{i=1}^{N}\mu_{i,t+h|t}$ denotes the equally weighted average (or 'consensus') forecast. Among others, this measure has been used by Lahiri and Sheng (2010), Glas and Hartmann (2016) and Abel et al. (2016). However, individual observations are relatively influential on the disagreement statistic in Eqn. (4.1) if N is small. To address this shortcoming, Abel et al. (2016) use the interquartile range (IQR) of the point forecasts,

$$\tilde{f}_{t+h|t} = \mu^{0.75}_{t+h|t} - \mu^{0.25}_{t+h|t}, \tag{4.2}$$

where $\mu^{0.25}_{t+h|t}$ and $\mu^{0.75}_{t+h|t}$ denote the 25% and 75% percentiles of the ordered array of point predictions, respectively.[1]

Several surveys of macroeconomic expectations such as the SPF elicit subjective probability distributions in the form of histograms in addition to the point forecasts. These h-step-ahead predictions consist of sequences of probabilities, $p_{i,k,t+h|t} \in [0,1]$ for $k = 1, \ldots, K$, where the index k indicates that probabilities are reported for a finite set of outcome intervals (so-called 'bins'). Each bin, $[l_k, u_k)$, is characterized by a lower bound, l_k, and an upper bound, u_k. Thus, $p_{i,k,t+h|t} = \Pr_i(x_{t+h} \in [l_k, u_k))$ denotes the probability that forecaster i assigns to the event that the outcome, x_{t+h}, falls into the range covered by the k-th bin. It is assumed that the probabilities sum to unity, i.e., $\sum_{k=1}^{K} p_{i,k,t+h|t} = 1$. Let

$$\mu^{\star}_{i,t+h|t} = \sum_{k=1}^{K} p_{i,k,t+h|t} \times m_k \tag{4.3}$$

denote the mean of forecaster i's histogram with $m_k = (l_k + u_k)/2$ indicating the midpoint of the k-th bin. In Eqn. (4.3), it is assumed that the probability mass in each bin is

[1] Boero et al. (2015) instead use one half of the difference between the 16% and 84% percentiles.

centered at the midpoint. This approach has been used, among others, by Poncela and Senra (2017).

In the case of a quadratic loss function, the point forecast and the histogram mean from Eqn. (4.3) theoretically coincide. However, Clements (2010, 2012) shows that $\mu_{i,t+h|t}$ and $\mu^{\star}_{i,t+h|t}$ tend to deviate in practice. This may occur in cases where forecasters use asymmetric loss functions. In this case, it is not clear whether the point predictions refer to the mean, median or mode of the histograms.[2] Rounding errors may also be a source of the misalignment between the two series. Another explanation is offered by Clements (2010), who documents that participants in the Federal Reserve's SPF update their point forecasts more frequently than the histogram means. In order to analyze whether the deviation of $\mu_{i,t+h|t}$ and $\mu^{\star}_{i,t+h|t}$ affects the relationship between uncertainty and disagreement, we additionally calculate the disagreement statistics from Eqns. (4.1) and (4.2) using $\mu^{\star}_{i,t+h|t}$. These measures are denoted as $s^{\star}_{t+h|t}$ and $\tilde{f}^{\star}_{t+h|t}$, respectively.

4.2.2 Individual Uncertainty

The measures from Eqns. (4.1) and (4.2) indicate the extent to which the survey participants disagree about future outcomes. The validity of using disagreement as a proxy for uncertainty has been frequently debated in the literature. Relevant studies in this area include Zarnowitz and Lambros (1987), Giordani and Söderlind (2003), Lahiri and Sheng (2010), Rich and Tracy (2010), Boero et al. (2008, 2015), Dovern (2015), Abel et al. (2016) and Glas and Hartmann (2016), among others. Intuitively, it may be argued that an appropriate measure of uncertainty should not only reflect the heterogeneity of individual expectations, but also capture how confident forecasters are about their own predictions. To investigate the usefulness of disagreement as a proxy for uncertainty, the former is usually compared to an aggregated measure of the second moments from the histogram forecasts, e.g., the average variance, which provides a more direct assessment of uncertainty. The derivation of the second moments requires an assumption about the distribution of the probability mass within the bins. Under the 'mass-at-midpoint' assumption discussed above, the variance of forecaster i's histogram, which serves as an indicator of subjective uncertainty, is given by

$$\sigma^2_{i,t+h|t,\mathrm{M}} = \sum_{k=1}^{K} p_{i,k,t+h|t} \times (m_k - \mu^{\star}_{i,t+h|t})^2. \tag{4.4}$$

[2] In a special survey conducted by the ECB in 2013, 49% of the SPF participants stated that their point forecasts represent the histogram mean, 18% indicated that they report the median and 16% replied that they submit the mode. The remaining panelists did not calculate probability distributions (ECB, 2014b).

Zarnowitz and Lambros (1987) and Abel et al. (2016) instead assume uniformly distributed probabilities within each bin due to the fact that the underlying outcome variable is usually continuous (e.g., the inflation rate). Note that the means derived from both approaches are identical, such that this choice does not affect the measures of disagreement based on $\mu^{\star}_{i,t+h|t}$, i.e., $s^{\star}_{t+h|t}$ and $\tilde{f}^{\star}_{t+h|t}$. However, the variance derived under the assumption of uniformly distributed probabilities deviates from the one in Eqn. (4.4) and is computed as

$$\sigma^2_{i,t+h|t,\mathrm{U}} = \left[\sum_{k=1}^{K} p_{i,k,t+h|t} \times \left(\frac{u_k^3 - l_k^3}{3(u_k - l_k)}\right)\right] - \left[\sum_{k=1}^{K} p_{i,k,t+h|t} \times \left(\frac{u_k^2 - l_k^2}{2(u_k - l_k)}\right)\right]^2. \tag{4.5}$$

If a forecaster assigns 100% probability to a single bin, the variance from Eqn. (4.4) is equal to zero, whereas the variance from Eqn. (4.5) is positive in this case. Since the histogram mean is not affected by this choice, the link between uncertainty and disagreement differs mechanically for distinct choices regarding the distribution of the probability mass within the bins. However, D'Amico and Orphanides (2008) argue that the difference in the variances from Eqns. (4.4) and (4.5) is small in most cases. Formally, Boero et al. (2015) show that the latter exceeds the former by $(u_k - l_k)^2/12$. In most surveys the bin width, $u_k - l_k$, is relatively narrow, so that the difference between the variances is small as well.

Giordani and Söderlind (2003) document that variances derived from the reported histograms tend to be overestimated. A popular alternative is to fit a continuous distribution to each histogram in order to approximate the underlying true subjective distribution. The aim is to choose the parameters $\boldsymbol{\theta}_{i,t,h}$ of the alleged distribution, such as to

$$\min_{\boldsymbol{\theta}_{i,t,h}} \sum_{k=1}^{K} (F - P_{i,k,t+h|t})^2, \tag{4.6}$$

where F is the cumulative distribution function (cdf) of the theoretical distribution and $P_{i,k,t+h|t} = \sum_{\tau=1}^{k} p_{i,\tau,t+h|t}$ is the empirically observed cdf. Various choices of F have been discussed in the empirical literature. We follow Engelberg et al. (2009), Clements (2014) and Glas and Hartmann (2016) and fit generalized beta distributions by choosing the parameters $\boldsymbol{\theta}_{i,t,h} = (\tilde{l}_{i,t,h}, \tilde{u}_{i,t,h}, a_{i,t,h}, b_{i,t,h})'$ such that the expression in Eqn. (4.6) is minimized.[3] The support of each beta distribution, $[\tilde{l}_{i,t,h}, \tilde{u}_{i,t,h}]$, is defined as the lower bound of the smallest bin with nonzero probability mass and the upper bound of the highest

[3] D'Amico and Orphanides (2008) instead consider a Gamma distribution. Giordani and Söderlind (2003), Lahiri and Liu (2006), D'Amico and Orphanides (2008), Söderlind (2011) and Boero et al. (2015) fit normal distributions. The Bank of England uses two-piece normal distributions in their fan charts to indicate the uncertainty of their economic outlook. Krüger (2017) also uses two-piece normal distributions.

bin with nonzero probability mass. The shape parameters are denoted as $a_{i,t,h}$ and $b_{i,t,h}$. Mean and variance of the beta distribution are given by

$$\mu_{i,t+h|t,\mathrm{B}} = \frac{a_{i,t,h}}{a_{i,t,h} + b_{i,t,h}} \times (\tilde{u}_{i,t,h} - \tilde{l}_{i,t,h}) + \tilde{l}_{i,t,h} \tag{4.7}$$

and

$$\sigma^2_{i,t+h|t,\mathrm{B}} = \frac{a_{i,t,h} b_{i,t,h}}{(a_{i,t,h} + b_{i,t,h} + 1)(a_{i,t,h} + b_{i,t,h})^2} \times (\tilde{u}_{i,t,h} - \tilde{l}_{i,t,h})^2, \tag{4.8}$$

respectively. A drawback of this approach is that forecasters are required to assign nonzero probabilities to at least three bins. If two or less bins are used, it is common practice to fit triangular distributions. For further details, we refer the reader to Engelberg et al. (2009).

The means of the beta distributions from Eqn. (4.7) are also used to derive measures of forecaster disagreement based on Eqns. (4.1) and (4.2), i.e., $s^{\mathrm{B}}_{t+h|t}$ and $\tilde{f}^{\mathrm{B}}_{t+h|t}$. The superscript 'B' indicates that the disagreement statistics are derived from the means of the beta distributions. The variance-based measure, $s^{\mathrm{B}}_{t+h|t}$, is analyzed in Glas and Hartmann (2016).

4.2.3 Aggregate Uncertainty

Based on the individual variances described above, measures of aggregate uncertainty can be derived. Following Boero et al. (2008, 2015), a commonly used indicator of uncertainty is the root mean subjective variance (RMSV),

$$\overline{\sigma}^j_{t+h|t} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \sigma^2_{i,t+h|t,j}}, \tag{4.9}$$

where $j \in \{\mathrm{M}, \mathrm{U}, \mathrm{B}\}$ indicates whether the RMSV series is based on the variances from Eqns. (4.4), (4.5) or (4.8).[4]

Abel et al. (2016) derive an alternative measure of uncertainty based on the IQR of forecaster i's probability distribution, i.e.,

$$\tilde{p}_{i,t+h|t} = p^{0.75}_{i,t+h|t} - p^{0.25}_{i,t+h|t}, \tag{4.10}$$

where $p^{0.25}_{i,t+h|t}$ and $p^{0.75}_{i,t+h|t}$ denote the 25% and 75% percentiles, respectively. Under the

[4] In order to account for outliers and data errors, Söderlind (2011) suggests the cross-sectional trimmed mean of the individual standard deviations with 20% trimming from top and bottom.

'mass-at-midpoint' approach, the quartiles in Eqn. (4.10) are defined as the midpoint of the bin for which the empirical cdf, $P_{i,k,t+h|t}$, first exceeds the respective threshold. If a uniform distribution is assumed, the quantiles in Eqn. (4.10) are derived by means of linear interpolation. For the parametric approach, the quantile functions of the triangular and beta distributions are used.

Following Abel et al. (2016), we use the cross-sectional median across the ordered array of IQR statistics from Eqn. (4.10),

$$\tilde{\phi}^{j}_{t+h|t} = \tilde{p}^{0.5}_{t+h|t}, \tag{4.11}$$

as an alternative indicator of uncertainty. The median-based statistic in Eqn. (4.11) is considered for the uniform and beta distributions, i.e., for $j \in \{\mathrm{U}, \mathrm{B}\}$. In the case of the 'mass-at-midpoint' approach, we use the cross-sectional average of the IQR statistics, i.e.,

$$\tilde{\phi}^{\mathrm{M}}_{t+h|t} = \frac{1}{N}\sum_{i=1}^{N} \tilde{p}^{\mathrm{M}}_{i,t+h|t}. \tag{4.12}$$

We use the average instead of the median in Eqn. (4.12) because the lower and upper quartiles derived under the 'mass-at-midpoint' assumption frequently coincide for many of the survey participants. Thus, the median uncertainty from Eqn. (4.11) is equal to zero in many instances. Nonetheless, $\tilde{\phi}^{\mathrm{M}}_{t+h|t}$ is considered to be more robust to outliers than $\overline{\sigma}^{\mathrm{M}}_{t+h|t}$ because the former is based on the IQRs of the individual distributions instead of the variances.

4.2.4 The Link Between Uncertainty and Disagreement

In order to analyze the relationship between uncertainty and disagreement in a formal way, we first relate the variance-based indicators of uncertainty to the variance-based measures of disagreement based on the point forecasts, i.e., we consider models of the form

$$\overline{\sigma}^{j}_{t+h|t} = \alpha + \beta \times s_{t+h|t} + \varepsilon^{j}_{t+h|t}, \tag{4.13}$$

where $\overline{\sigma}^{j}_{t+h|t}$ for $j \in \{\mathrm{M}, \mathrm{U}, \mathrm{B}\}$ is the RMSV statistic defined in Eqn. (4.9) and $s_{t+h|t}$ denotes the standard deviation of the point predictions from Eqn. (4.1). This is the most commonly used approach in the empirical literature (e.g., Lahiri and Sheng, 2010).

In a second step, we asses the importance of outliers and analyze the link between the

IQR-based measures of disagreement and uncertainty via

$$\tilde{\phi}^{j}_{t+h|t} = \alpha + \beta \times \tilde{f}_{t+h|t} + \tilde{\varepsilon}^{j}_{t+h|t}, \tag{4.14}$$

where $\tilde{\phi}^{j}_{t+h|t}$ with $j \in \{\mathrm{M}, \mathrm{U}, \mathrm{B}\}$ is one of the IQR-based measures of uncertainty from Eqn. (4.11) and (4.12) and $\tilde{f}_{t+h|t}$ denotes the IQR of the point forecasts from Eqn. (4.2).

In the third step, we account for the well-documented difference between the point forecasts and the histogram means. To this end, we replace the disagreement statistics based on the point forecasts in Eqns. (4.13) and (4.14) with the corresponding measures derived from the histogram means, i.e., $s^{\star}_{t+h|t}$ and $\tilde{f}^{\star}_{t+h|t}$ in the cases of the 'mass-at-midpoint' and uniform assumptions and $s^{\mathrm{B}}_{t+h|t}$ and $\tilde{f}^{\mathrm{B}}_{t+h|t}$ for the beta distribution.

For each combination of outcome variable and forecast horizon, the three-step procedure described above yields twelve specifications of Eqns. (4.13) and (4.14) in total. In each case, disagreement is a perfect proxy for uncertainty if $\alpha = 0$ and $\beta = 1$. A more conservative approach is to allow for a mean bias and to test whether $\beta = 1$. A minimal requirement is that the two series are correlated, i.e., $\beta \neq 0$.

Our analysis serves as an extension of the analysis in Abel et al. (2016), who account for different dispersion statistics, outcome variables and forecast horizons but do not consider alternative assumptions about the distribution of the probabilities within the bins and disregard the difference between the point forecasts and the means of the individual distributions. In other words, Abel et al. (2016) analyze two out of the twelve specifications discussed above by relating $\overline{\sigma}^{\mathrm{U}}_{t+h|t}$ to $s_{t+h|t}$ and $\tilde{\phi}^{\mathrm{U}}_{t+h|t}$ to $\tilde{f}_{t+h|t}$. The relationship between $\overline{\sigma}^{\mathrm{B}}_{t+h|t}$ and $s^{\mathrm{B}}_{t+h|t}$ is analyzed in Glas and Hartmann (2016).

4.3 Data

This section describes the survey data that are used to quantify the measures of disagreement and uncertainty. The data are taken from the ECB's SPF, which provides individual forecasts of the harmonized index of consumer prices inflation, real GDP growth, and unemployment rates in the Euro area for a heterogeneous cross-section of expert forecasters including employees from both the financial sector and research institutions. The SPF has been conducted in consecutive quarters since 1999Q1 and elicits both point and histogram forecasts. Participation in the survey is anonymous but panelists can be traced over time with the help of an identification number that is assigned to each individual response. In the empirical analysis, we use the one-year- and two-years-ahead fixed-horizon forecasts as well as the long-term fixed-event forecasts from the surveys conducted between 1999Q1 and 2018Q2, such that $T = 78$.[5] Our sample thus exceeds the one used by Abel et al. (2016) by more than four years and covers politically relevant events such as the decision of the UK to leave the EU after the outcome of the referendum in 2016Q2. As noted by Abel et al. (2016) and Rossi and Sekhposyan (2017), the actual forecast horizons differ across variables and survey periods due to differences in data frequencies and publication lags. Moreover, the horizon of the long-term predictions declines from 22- to 19-quarters-ahead for the Q1–Q4 surveys because the fixed-event forecasts refer to annual outcomes. To improve the readability, we nonetheless denote the forecast horizons as $h \in \{4, 8, 20\}$, respectively.

It is important to acknowledge certain features of the histograms provided by the SPF. First, the interior bins have a width of 0.4 percentage points with a gap of 0.1 percentage points between bins, e.g., $[1.0, 1.4)$ and $[1.5, 1.9)$. Following Abel et al. (2016), this gap is closed by extending the lower and upper bound of each bin by 0.05 percentage points, e.g., $[0.95, 1.45)$ and $[1.45, 1.95)$.[6] Second, the half-open exterior bins are closed by assuming that they are of the same length as the interior intervals, i.e., 0.5 percentage points.[7] Third, the bounds of each histogram are fixed at the left-most and right-most bin with nonzero probability mass. Fourth, the range of available bins for each outcome variable changes over time. Such adjustments may be the result of an observed pile-up of probabilities in one of the exterior bins. A particularly noteworthy occurence is discussed below.

In order to obtain a homogeneous data set and reduce the influence of outliers, we exclude the following observations from the sample: First, we follow Rich and Tracy (2010) and Abel et al. (2016) and consider matched point and density forecasts. This means that forecasters are required to report both point *and* histogram forecasts at a given survey date for a combination of outcome variable and forecast horizon.[8] This procedure ensures that in each quarter, disagreement and uncertainty are calculated based on forecasts from the same group of SPF participants. Second, observations are excluded from the sample whenever the reported probabilities do not sum to unity. More precisely, we follow Abel et

[5] The long-term forecasts have not been elicited for the surveys conducted during the periods 1999Q2-1999Q4 and 2000Q2-2000Q4.

[6] In a special survey conducted by the ECB in 2008, 76% of the SPF participants stated that they interpret an interval such as $[1.5, 1.9)$ to actually indicate $[1.45, 1.95)$ (cf. ECB, 2009).

[7] The same approach is used by Zarnowitz and Lambros (1987), Lahiri and Liu (2006) and D'Amico and Orphanides (2008). Abel et al. (2016) instead assume that the exterior bins have twice the length of the interior bins. Our findings are robust to this choice. Alternative results will be provided upon request.

[8] However, forecasters are not required to report matched forecasts for *each* variable and/or forecast horizon. Thus, it is permitted that in a particular survey period a forecaster reports point and histogram forecasts for next year's inflation rate, but neither for next year's output growth, nor the inflation rate two years from now.

al. (2016) and delete all responses for which the total probability mass deviates by at least 0.9% from 100% in absolute terms. In Table 4.1, we report the number and share of total, valid and invalid responses for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$.

Table 4.1: Valid and invalid responses in the SPF data

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
Total	4476	4476	4123	4476	4476	4124	4476	4477	4124
Valid	3420	3084	2769	3327	3050	2735	3069	2809	2494
Share valid	76%	69%	67%	74%	68%	66%	69%	63%	60%
Invalid	1056	1392	1354	1149	1426	1389	1407	1668	1630
Share invalid	24%	31%	33%	26%	32%	34%	31%	37%	40%

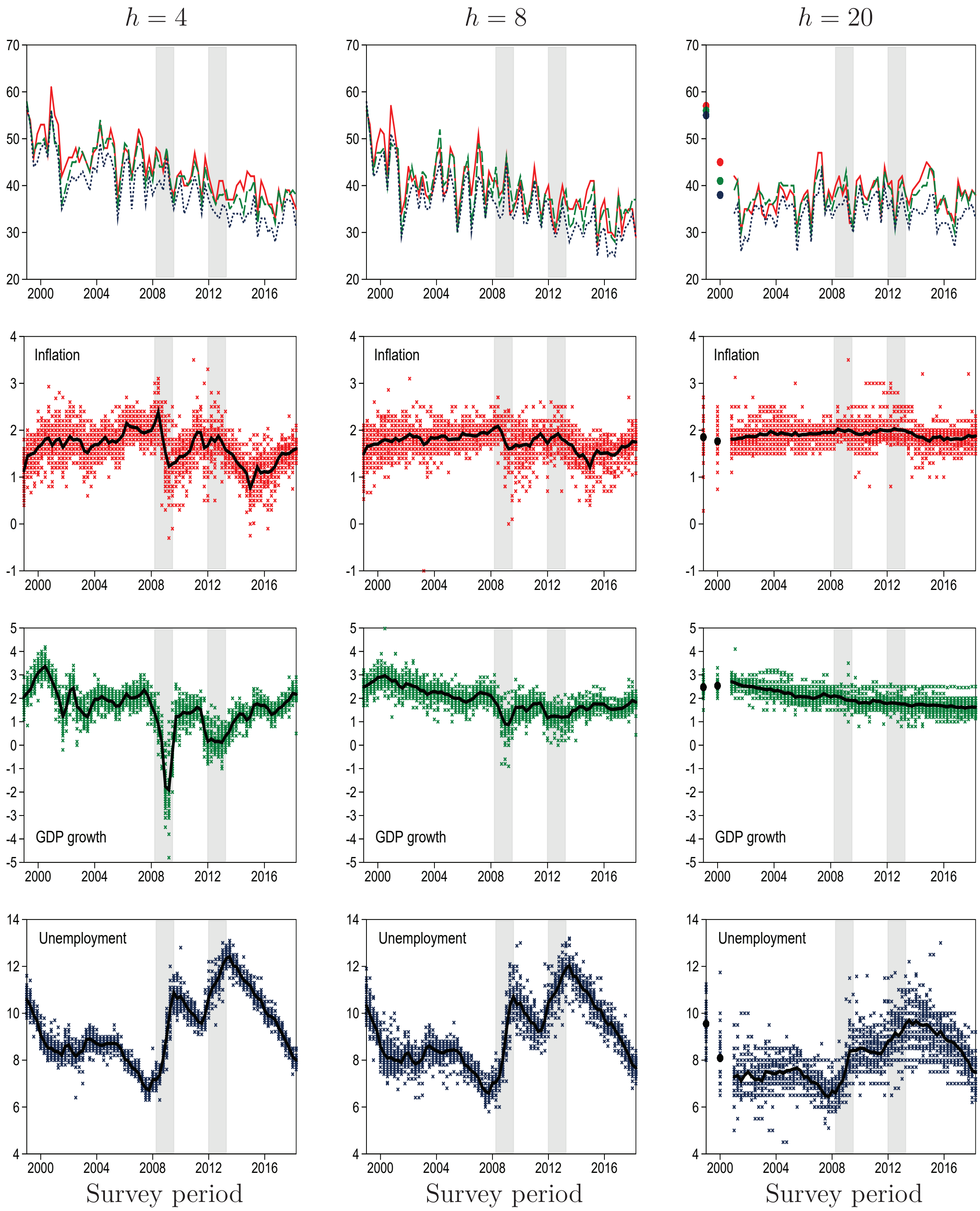
Notes: This table displays the total number of SPF responses as well as the number (and share) of valid and invalid responses for the sample period 1999Q1–2018Q2. Responses are considered to be invalid if forecasters do not report matched point and histogram forecasts or whenever the probabilities do not sum to unity.

On average, a relatively large fraction of 32% of the responses are excluded from the sample. Out of those cases, 68% represent instances where neither a point nor a density forecast is provided, 29% of these predictions do not include a histogram and 3% are missing the point forecast. In contrast, only three observations are excluded because the reported probabilities do not sum to 100%. Notably, the share of invalid responses increases with the forecast horizon. This may be an indication that forecasters put less effort into reporting adequate long-term predictions.

The first row of Figure 4.1 depicts the number of valid SPF responses in each survey round. Shaded gray areas indicate recession periods as classified by the Centre for Economic Policy Research (CEPR). So far, we have implicitly assumed that the survey panel is balanced, i.e., all forecasters submit predictions in consecutive quarters. Figure 4.1 shows that participation in the SPF has declined considerably over time. Depending on the outcome variable and forecast horizon, 55–58 valid responses have been reported in 1999Q1, whereas the number in 2018Q2 varies between 29 and 38. However, the number of predictions per survey round appears to have stabilized in recent years. Moreover, participation in the SPF fluctuates across quarters. This seasonal pattern has also been observed by López-Pérez (2016) and is visible more clearly in Table 4.2, which displays the average number of responses for each of the four quarters.

For each outcome variable and forecast horizon, the average participation in the SPF is highest in the Q1-surveys and lowest in the Q3-surveys. This suggests that the cross-section of forecasters may systematically differ across quarters, e.g., due to a summer

Figure 4.1: Participation and point forecasts in the ECB-SPF



Notes: The plots in the first row depict the number of valid SPF responses per survey round for the inflation (solid line), output growth (dashed line), and unemployment rates (dotted line) at forecast horizons $h \in \{4, 8, 20\}$. The remaining rows depict the individual point predictions, $\mu_{i,t+h|t}$ (crosses '×'), and consensus forecasts, $\overline{\mu}_{t+h|t}$ (solid line), for the inflation (second row), output growth (third row), and unemployment rates (fourth row) at each forecast horizon. The horizontal axis depicts the period during which predictions are reported and covers the period 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

Table 4.2: Average participation in the ECB-SPF

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
Q1-surveys	46.49	44.50	41.20	45.04	42.98	40.63	42.01	40.48	36.82
Q2-surveys	45.10	40.39	38.87	44.49	42.89	38.91	40.82	37.51	35.51
Q3-surveys	41.48	36.02	35.82	39.94	34.46	35.08	36.98	32.72	32.27
Q4-surveys	44.89	39.98	38.90	43.62	38.29	38.04	40.25	36.06	34.75

Notes: For each outcome variable and forecast horizon, this table displays the average number of valid SPF responses per quarter. The sample period is 1999Q1–2018Q2.

break in certain institutions. Taken together, Figure 4.1 and Table 4.2 provide substantial evidence of entry and exit of forecasters to and from the SPF panel. Capistran and Timmermann (2009) show that forecast combinations are biased in this case. In order to address this issue, we present the results of the empirical analysis based on both the entire cross-section and a smaller sample of the most regular SPF participants.

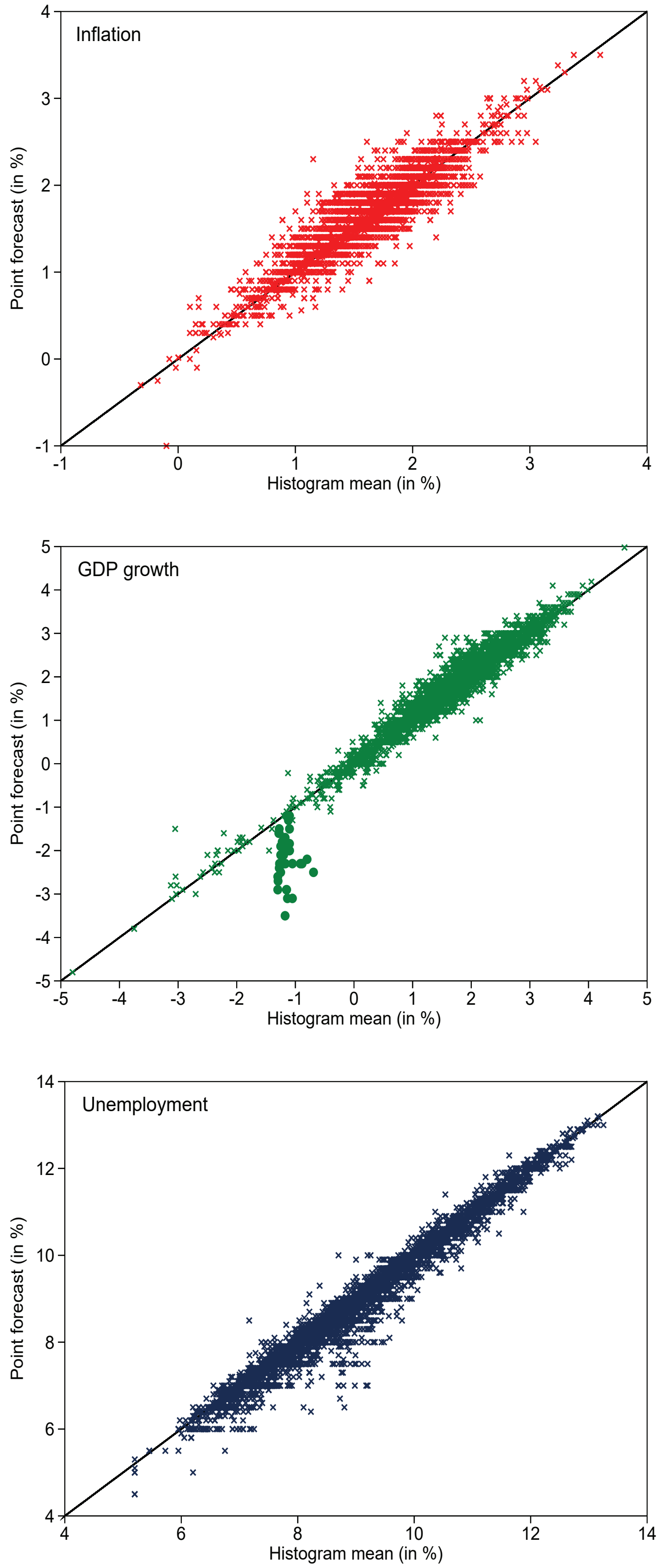
The plots in the remaining rows of Figure 4.1 depict the individual point predictions, $\mu_{i,t+h|t}$, and the corresponding consensus forecast, $\overline{\mu}_{t+h|t}$, for each outcome variable. The one-year-ahead forecasts of the inflation and output growth rates fell sharply after the outbreak of the Great Recession. In contrast, the long-term predictions remain relatively stable throughout the sample period. As a result, the cross-sectional variation of both series appears to decline in the forecast horizon. Consistent with the expected decline in economic activity, the predicted unemployment rate rose considerably around both 2008 and 2012. In contrast to inflation and GDP growth, the cross-sectional dispersion of the expected unemployment rate increases with h.

As mentioned in Section 4.2, a misalignment of the point forecasts and the histogram means has been frequently observed (Clements, 2010, 2012). In order to analyze how strongly the two series deviate in the SPF data, Figure 4.2 depicts the individual point predictions, $\mu_{i,t+h|t}$, on the vertical axis and the histogram means, $\mu^{\star}_{i,t+h|t}$, on the horizontal axis for a pooled sample of observations across all forecasters, time periods, and forecast horizons.[9]

Overall, the two series are very similar, but not identical. The correlation statistics between $\mu_{i,t+h|t}$ and $\mu^{\star}_{i,t+h|t}$ in the pooled sample for the inflation, output growth, and unemployment rates are equal to 0.91, 0.97, and 0.99, respectively. Notably, there is a cluster of predictions for real GDP growth with point predictions that are considerably

[9] The means of the histograms and the beta distributions, i.e., $\mu^{\star}_{i,t+h|t}$ and $\mu^{\mathrm{B}}_{i,t+h|t}$, are almost identical. Thus, we do not discuss the deviations of $\mu_{i,t+h|t}$ and $\mu^{\mathrm{B}}_{i,t+h|t}$ here.

Figure 4.2: Point forecasts versus histogram means



Notes: For a pooled sample of observations across all forecasters, time periods and forecast horizons $h \in \{4, 8, 20\}$, this figurethe plots depict the individual point forecasts, $\mu_{i,t+h|t}$ on the vertical axis and the histogram means, $\mu^{\star}_{i,t+h|t}$ on the horizontal axis for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel). Bullets '•' in the middle panel indicate notable predictions of real GDP from the 2009Q1 survey. The sample period is 1999Q1–2018Q2.

lower than the histogram means near the bottom-left corner of the plot (highlighted as bullets '•'). These observations were submitted in the 2009Q1 survey, for which a substantial pile-up of probabilities in the lowest bin has been observed by Abel et al. (2016), among others. In reaction to this, the bin definitions were adjusted to allow for a wider range of possible outcomes. As a result of the pile-up, the reported histogram means from this survey are substantially larger than the point forecasts in many cases and also less dispersed. This conglomeration induces a downward bias in the disagreement statistics based on the means and also affects the uncertainty measures. In order to avoid potentially midleading conclusions, we exclude the output growth forecasts from the 2009Q1 survey from the empirical analysis in Section 4.4. More generally, the observed deviations of the series in Figure 4.2 justify the use of the disagreement statistics based on both the point forecasts and the histogram means.

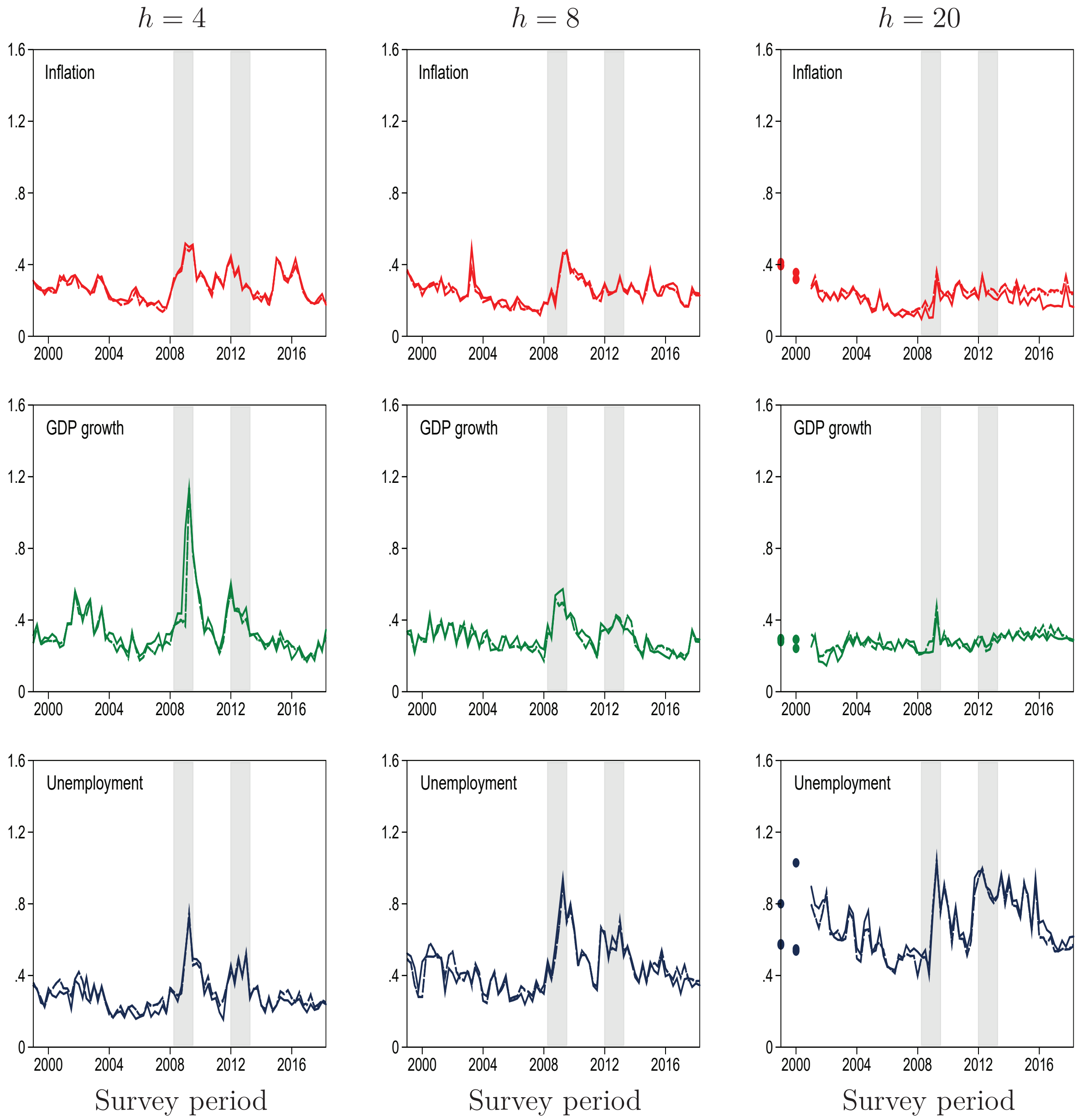
In the next step, we describe the time series of the disagreement statistics. Figures 4.3 and 4.4 depict the evolution of the variance- and IQR-based measures of forecaster disagreement from Eqns. (4.1) and (4.2), respectively. Each plot contains the series based on the point forecasts (solid line), the histogram means (dashed line), and the means of the beta distributions (dotted line).

All disagreement measures evolve in a relatively similar way. In line with Andrade et al. (2016), disagreement is nonzero and time-varying for each outcome variable and forecast horizon. In most cases, surges in disagreement, such as during the Great Recession and the sovereign debt crisis, are followed by a relatively quick return to low levels. Variance-based disagreement is generally lower than the IQR-based series. Moreover, the differences in the statistics based on the point forecasts and the means of either the histograms or the beta distributions are negligible in Figure 4.3, but not in Figure 4.4.

In line with the impression from Figure 4.1, disagreement for both the inflation rate and output growth declines in the forecast horizon, whereas the opposite is the case for the unemployment rate. Andrade et al. (2016) document a similar upward-sloping term structure of forecaster disagreement for the expectations of the federal funds rate from the Blue Chip Financial Forecasts survey. This finding is explained by the fact that the federal funds rate is highly persistent, and thus easy to predict in the short-term using its most recent value. A similar argument can be used in the case of the unemployment rate, which evolves more smoothly than both the inflation rate and economic growth. This may explain the deviations of the disagreement statistics across the three outcome variables.

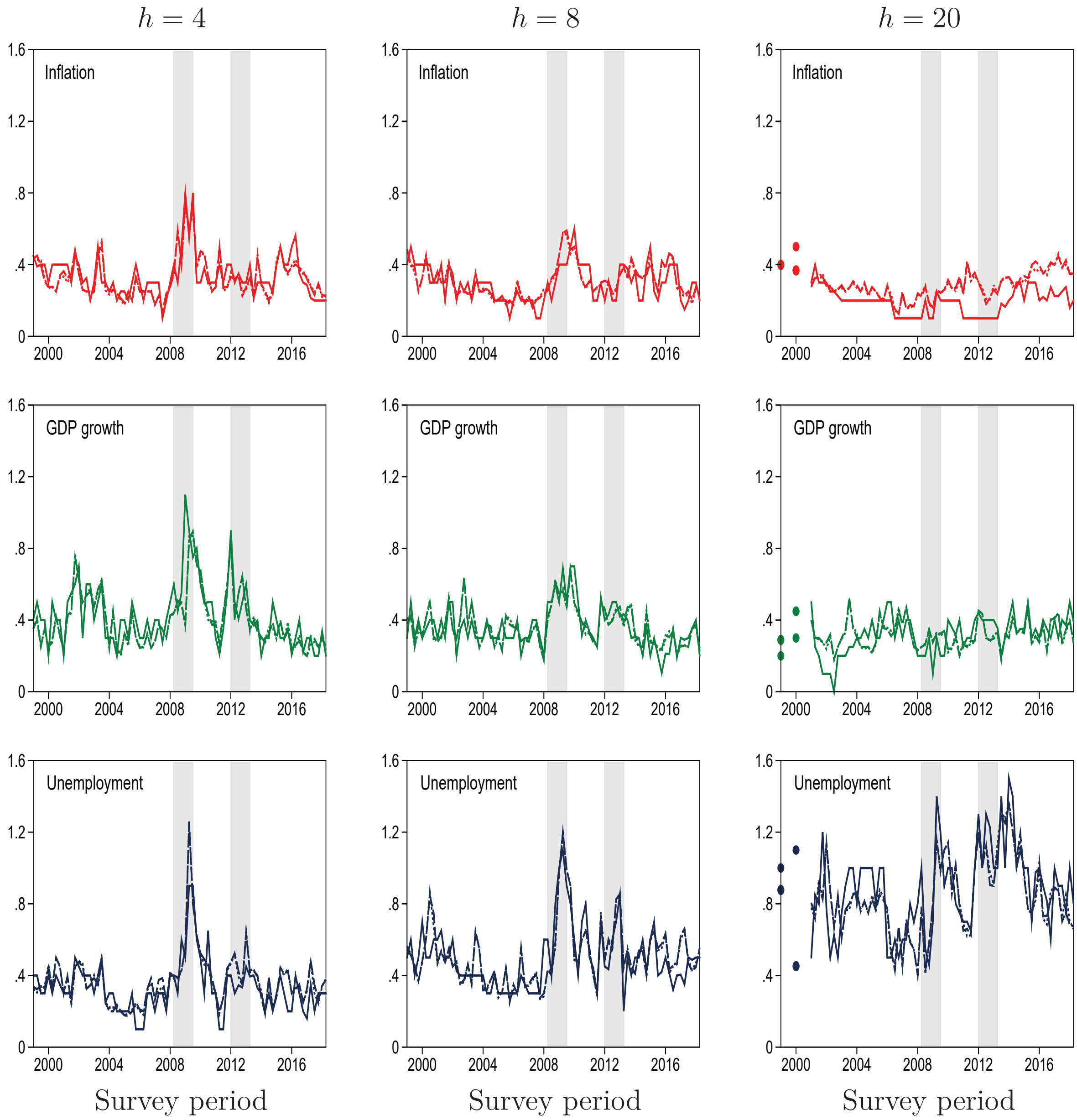
The observed behavior of the disagreement series is notably different from the evolution of the RMSV- and IQR-based measures of uncertainty from Eqns. (4.9), (4.11), and (4.12),

Figure 4.3: Variance-based disagreement



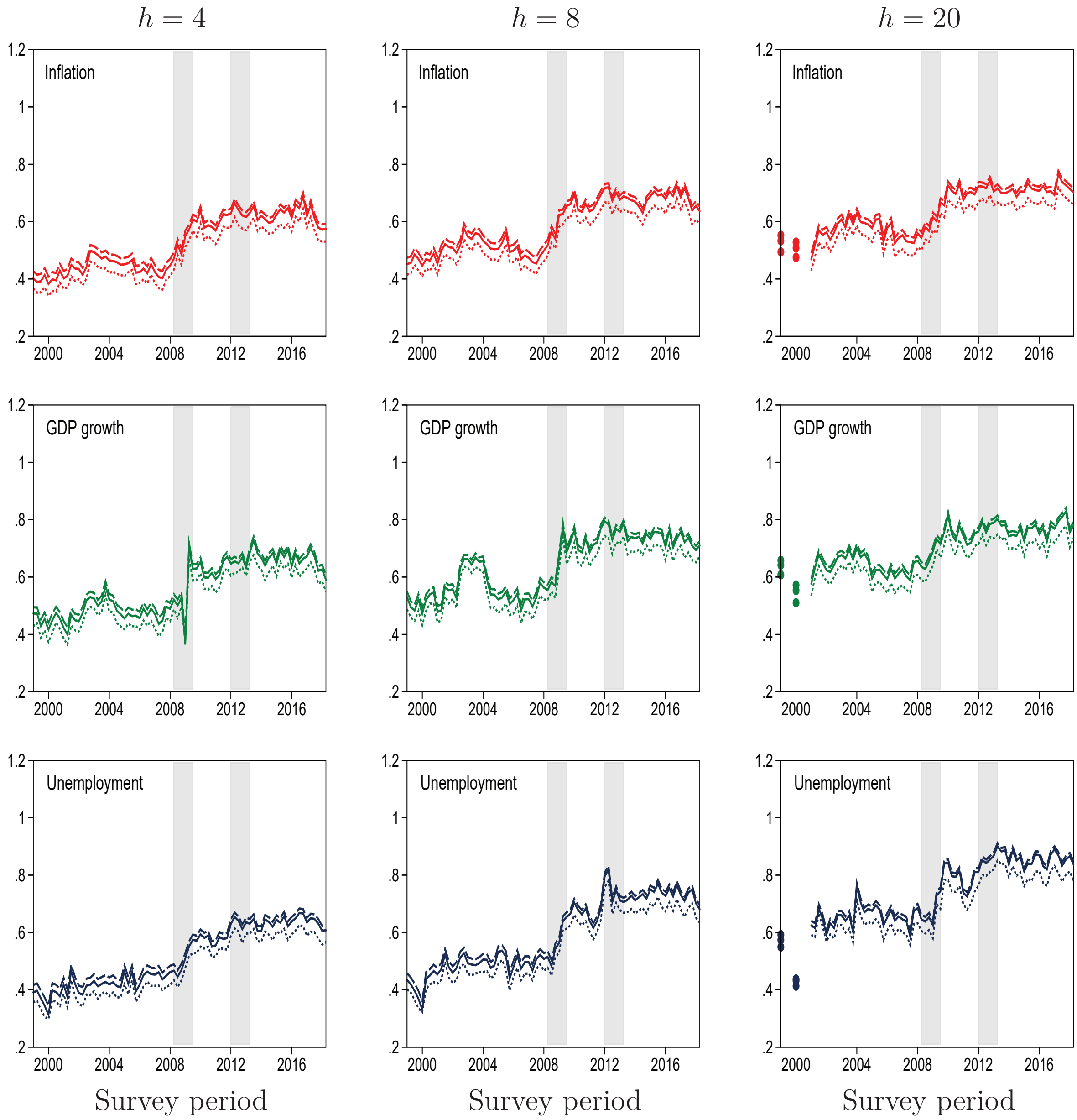
Notes: These plots depict the evolution of the variance-based disagreement statistic from Eqn. (4.1) for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel) at forecast horizons $h \in \{4, 8, 20\}$. Each figure includes the disagreement series based on the point forecasts (solid line), the histogram means (dashed line), and the means of the beta distributions (dotted line), i.e., $s_{t+h|t}$, $s^{\star}_{t+h|t}$, and $s^{\mathrm{B}}_{t+h|t}$, respectively. The horizontal axis depicts the period during which predictions are reported and covers the surveys from 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

Figure 4.4: IQR-based disagreement



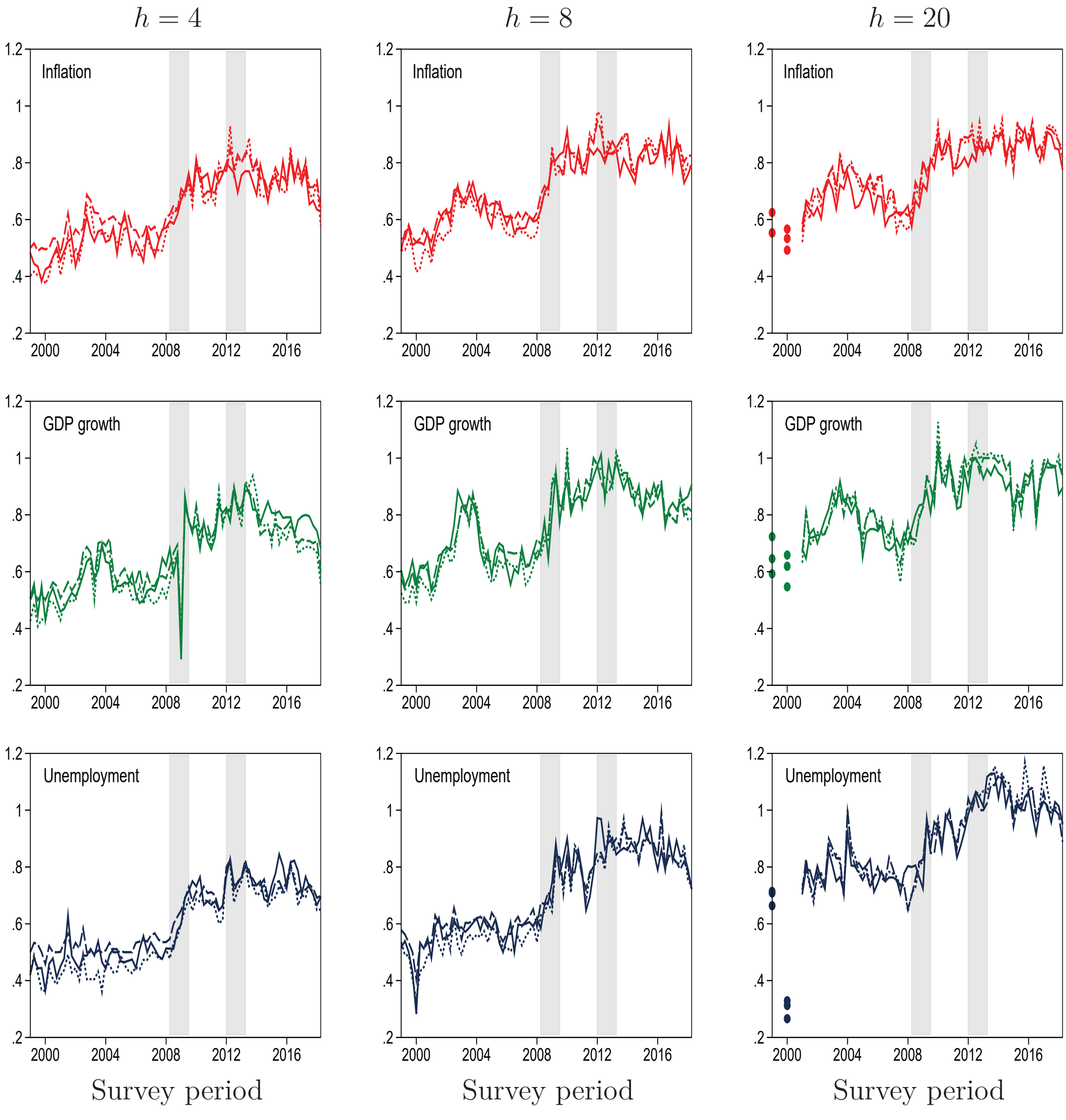
Notes: These plots depict the evolution of the IQR-based disagreement statistic from Eqn. (4.2) for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel) at forecast horizons $h \in \{4, 8, 20\}$. Each figure includes the disagreement series based on the point forecasts (solid line), the histogram means (dashed line), and the means of the beta distributions (dotted line), i.e., $\tilde{f}_{t+h|t}$, $\tilde{f}^{\star}_{t+h|t}$, and $\tilde{f}^{\mathrm{B}}_{t+h|t}$, respectively. The horizontal axis depicts the period during which predictions are reported and covers the surveys from 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

Figure 4.5: RMSV-based uncertainty



Notes: These plots depict the evolution of the RMSV uncertainty statistic from Eqn. (4.9) for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel) at forecast horizons $h \in \{4, 8, 20\}$. Each figure includes the RMSV series based on the 'mass-at-midpoint' assumption (solid line), uniformly distributed probabilities (dashed line), and the fitted beta distributions (dotted line), i.e., $\overline{\sigma}^{\mathrm{M}}_{t+h|t}$, $\overline{\sigma}^{\mathrm{U}}_{t+h|t}$, and $\overline{\sigma}^{\mathrm{B}}_{t+h|t}$, respectively. The horizontal axis depicts the period during which predictions are reported and covers the surveys from 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

Figure 4.6: IQR-based uncertainty



Notes: These plots depict the evolution of the IQR-based uncertainty statistics from Eqns. (4.11) and (4.12) for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel) at forecast horizons $h \in \{4, 8, 20\}$. Each figure includes the IQR-based uncertainty series based on the 'mass-at-midpoint' assumption (solid line), uniformly distributed probabilities (dashed line), and the fitted beta distributions (dotted line), i.e., $\tilde{\phi}^{\text{M}}_{t+h|t}$, $\tilde{\phi}^{\text{U}}_{t+h|t}$, and $\tilde{\phi}^{\text{B}}_{t+h|t}$, respectively. The horizontal axis depicts the period during which predictions are reported and covers the surveys from 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

which are depicted in Figures 4.5 and 4.6, respectively. Each plot contains the time series of uncertainty based on the 'mass-at-midpoint' approach (solid line), the assumption of uniformly distributed probabilities (dashed line), and the beta distributions (dotted line).

Across all statistics, outcome variables, and forecast horizons, an increase in aggregate uncertainty is visible after the outbreak of the Great Recession in 2008. Unlike the disagreement series, uncertainty does not return to its pre-crisis level. Rather, all measures remain elevated throughout the rest of the sample.

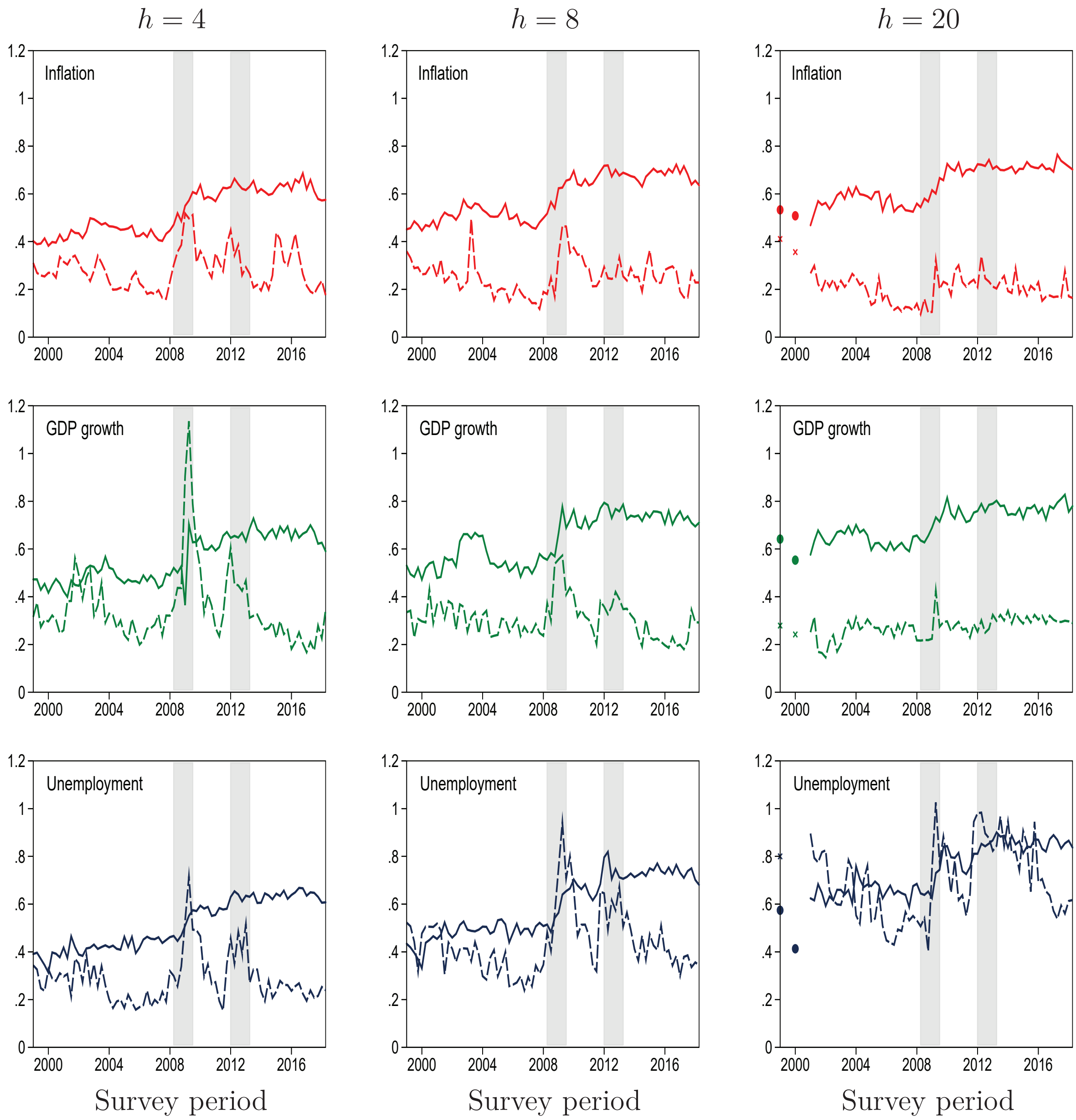
With respect to the different distributional assumptions, the RMSV series in Figure 4.5 evolve almost identically, except for small level shift. This is consistent with Boero et al. (2015), who note that measures based on the assumption of uniformly distributed probabilities exceed those under the 'mass-at-midpoint' approach by approximately 0.02 percentage points based on the bin width of 0.5 percentage points. Moreover, Giordani and Söderlind (2003) and D'Amico and Orphanides (2008) note that variances derived from parametric distributions are smaller than the variances of the histograms. This is in line with the fact that the series based on the beta and triangular distributions are smaller than ones based on the histograms.[10] In sum, the patterns observed in Figure 4.5 reveal that $\overline{\sigma}^{\mathrm{U}}_{t+h|t} > \overline{\sigma}^{\mathrm{M}}_{t+h|t} > \overline{\sigma}^{\mathrm{B}}_{t+h|t}$. However, this distribution-based ranking is not apparent for the IQR-based series in Figure 4.6. Both figures illustrate the impact of the pile-up of probabilities in the lowest bin for the one-year-ahead output growth forecasts as well as a subsequent sharp increase in uncertainty after the adjustment of the bin definitions in 2009Q2.

López-Pérez (2016) shows that participation in the SPF is negatively related to the overall level of uncertainty. This observation squares with the patterns depicted in Figures 4.1, 4.5, and 4.6 and suggests that some of the most uncertain respondents may choose to leave the survey. Moreover, Clements (2014) finds that forecasters in the FED-SPF are 'overconfident' at forecast horizons of one year or more, i.e., the reported ex-ante variances are frequently too small compared to the ex-post variances derived from the time series of forecast errors. Similar findings are reported by Giordani and Söderlind (2003, 2006) for the FED-SPF as well as Kenny et al. (2015) and Krüger (2017) for the ECB-SPF. Thus, the observed level of uncertainty may only represent a lower bound of the true level.

As a consequence of the evolution of the measures depicted in Figures 4.3 to 4.6, the gap between disagreement and uncertainty has been increasing since the outbreak of the Great Recession across all outcome variables. However, the magnitude of the

[10] Across all forecast horizons, the share of SPF forecasters using fewer than three bins is 9–10% depending on the outcome variable. Thus, there are not many cases in which triangular distributions are used.

Figure 4.7: Decoupling of uncertainty and disagreement



Notes: These plots depict the evolution of the RMSV uncertainty statistic from Eqn. (4.9) based on the 'mass-at-midpoint' assumption, i.e., $\overline{\sigma}^{\mathrm{M}}_{t+h|t}$ (solid line), and the variance-based disagreement statistic from Eqn. (4.1) based on the point forecasts, i.e., $s_{t+h|t}$ (dashed line), for the inflation (upper panel), output growth (middle panel), and unemployment rates (lower panel) at forecast horizons $h \in \{4, 8, 20\}$. The horizontal axis depicts the period during which predictions are reported and covers the surveys from 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

deviation of the two series varies across variables. An example of this is provided in Figure 4.7, which shows the RMSV uncertainty statistic based on the 'mass-at-midpoint' assumption, i.e., $\overline{\sigma}^{\mathrm{M}}_{t+h|t}$, and the variance-based disagreement statistic based on the point forecasts, i.e., $s_{t+h|t}$.

The evolution of the series depicted in Figure 4.7 suggests that the strength of the link between disagreement and uncertainty may vary for distinct outcome variables. Moreover, Figures 4.3 to 4.6 show that both disagreement and uncertainty vary with the employed dispersion statistic, distributional assumptions and the forecast horizon. Based on a theoretical forecast error decomposition, Lahiri and Sheng (2010) show that the difference between both measures increases with the forecast horizon due to an accumulation of unpredictable aggregate shocks. Glas and Hartmann (2016) rely on the same decomposition to show that the two measures deviate more strongly during periods when monetary policy is unduly expansive. We conduct a formal analysis of the dimensions that affect the relationship between uncertainty and disagreement below.

4.4 The Uncertainty-Disagreement Linkage

In this section, we present the estimates of the empirical models from Eqns. (4.13) and (4.14). By estimating each each of the twelve specifications for the three outcome variables (inflation, output growth, unemployment) and forecast horizons $h \in \{4, 8, 20\}$, we obtain a set of $12 \times 3 \times 3 = 108$ estimates in total. To our knowledge, this is the most comprehensive evaluation of the link between uncertainty and disagreement in surveys of macroeconomic expectations. In contrast, Abel et al. (2016) only consider 18 out of these 108 specifications. In each case, the parameter vector $(\alpha, \beta)'$ is estimated via ordinary least squares (OLS). In order to account for the autocorrelation patterns in the data induced by the overlapping forecast horizons, we use the variance-covariance estimator proposed by Newey and West (1987).

The results are summarized in Table 4.3. In the first row, we report the average estimates of α and β across all 108 specifications $(\bar{\hat{\alpha}}, \bar{\hat{\beta}})'$ as well as the average goodness of fit ($\overline{R^2}$). For each estimate, we test the hypotheses $H_0 : \beta = 0$ and $H_0 : \beta = 1$ and report the percentage of cases in which the respective null hypothesis is rejected at the 5% critical level.[11] Abel et al. (2016) use one-sided tests to assess whether the estimated slope coefficients are significantly different from zero. However, in a few instances we obtain *negative* estimates of β. Therefore, the conventional two-sided alternative is considered to be more appropriate. The remaining rows of Table 4.3 describe the results across the five previously discussed dimensions of interest. These findings are obtained by conditioning the estimates on either the employed dispersion statistic (variance- vs. IQR-based), the

[11] We have also tested the joint hypothesis that $\alpha = 0$ and $\beta = 1$. However, the null hypothesis was rejected for all 108 specifications. Thus, we do not report these results in Table 4.3.

considered disagreement measure (point forecasts vs. histogram means), the assumed distribution ('mass-at-midpoint', uniform, beta), the outcome variable (inflation, output growth, unemployment) or the forecast horizon ($h \in \{4, 8, 20\}$).

In addition to the results based on the full sample of observations in columns 3–7, we report the results from two robustness checks. To evaluate the impact of the Great Recession and the subsequent expansionary monetary policy of the ECB, we estimate Eqns. (4.13) and (4.14) on a subsample of observations from 2008Q2 onward (columns 8–12). As a means to account for the frequent entry and exit of SPF participants, we additionally present the findings when only the most regular SPF participants are considered in the calculation of the disagreement and uncertainty measures from Section 4.2 (columns 13–17). A detailed description is provided below.

In the discussion of the results we will focus on the summary in Table 4.3 and only refer to individual estimates when necessary. In the Appendix we provide a replication of the results of Abel et al. (2016) (Table 4.5) as well as the individual estimates that underlie the findings from Table 4.3 (Tables 4.6–4.11).

The models underlying the estimates from Panels B and E of Table 4.6 correspond to the specifications considered in Abel et al. (2016). When we estimated these models on their sample period, i.e., 1999Q1–2013Q4, we obtained very similar results to the ones from their Table II (see Table 4.5). Our analysis is based on a larger sample of observations that includes four and a half years of additional data from the SPF. In almost all cases, the findings from Panels B and E of Table 4.6 are weaker than the evidence documented in Abel et al. (2016) in terms of both the magnitude and the statistical significance of the estimated coefficients. This serves to illustrate the political and economic importance of the most recent years in the sample, during which the UK decided to leave the EU (2016Q2) and Donald Trump was elected as the president of the United States (2017Q1).

4.4.1 Full-Sample Estimates

In the first step of the analysis, we estimate the 108 specifications based on Eqns. (4.13) and (4.14) for the full sample of SPF surveys, covering the period 1999Q1–2018Q2. The average estimates of the intercept and the slope coefficient reported in the first row of Table 4.3 are 0.57 and 0.25, respectively. Thus, the estimates differ considerably from the case where $\alpha = 0$ and $\beta = 1$. Out of the 108 specifications, we obtain 37 significant estimates of β, i.e., approximately 34% (see Tables 4.6–4.7). All of them are positive. However, the null hypothesis that $\beta = 1$ is rejected in 89% of all cases. On average, only 8% of the variation in the uncertainty series is explained by the movement in the disagreement statistics. Thus, the unconditional results provide little evidence for the

Table 4.3: Strength of the link between uncertainty and disagreement in the SPF data

		Full-sample estimates					Sample split					Regular participants				
	Models	$\bar{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\beta = 0$	$\beta = 1$	$\overline{R^2}$	$\bar{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\beta = 0$	$\beta = 1$	$\overline{R^2}$	$\bar{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\beta = 0$	$\beta = 1$	$\overline{R^2}$
Overall	108	0.57	0.25	34%	89%	0.08	0.71	0.10	19%	95%	0.07	0.60	0.21	22%	88%	0.06
Dispersion statistic																
Variance-based	54	0.51	0.29	35%	83%	0.08	0.65	0.08	11%	94%	0.06	0.54	0.21	28%	85%	0.06
IQR-based	54	0.64	0.21	33%	94%	0.07	0.78	0.12	26%	96%	0.08	0.66	0.20	17%	91%	0.05
Disagreement measure																
Point forecasts	54	0.59	0.19	22%	94%	0.05	0.71	0.11	19%	96%	0.08	0.62	0.15	17%	93%	0.05
Histogram means	54	0.55	0.30	46%	83%	0.10	0.72	0.09	19%	94%	0.06	0.58	0.26	28%	83%	0.06
Distributional assumption																
'Mass-at-midpoint'	36	0.57	0.26	33%	89%	0.08	0.72	0.10	22%	97%	0.07	0.60	0.21	28%	89%	0.06
Uniform distribution	36	0.60	0.24	33%	89%	0.08	0.73	0.09	19%	94%	0.07	0.62	0.20	19%	89%	0.05
Beta distribution	36	0.55	0.25	36%	89%	0.08	0.70	0.10	14%	94%	0.07	0.58	0.21	19%	86%	0.06
Outcome variable																
Inflation rate	36	0.56	0.28	28%	83%	0.07	0.69	0.13	14%	92%	0.06	0.58	0.21	22%	86%	0.05
GDP growth rate	36	0.62	0.23	22%	83%	0.06	0.73	0.14	25%	94%	0.07	0.61	0.27	25%	78%	0.07
Unemployment rate	36	0.53	0.24	53%	100%	0.11	0.72	0.03	17%	100%	0.08	0.60	0.14	19%	100%	0.04
Forecast horizon																
One-year-ahead	36	0.52	0.18	28%	100%	0.04	0.68	−0.02	3%	100%	0.02	0.54	0.12	19%	100%	0.02
Two-years-ahead	36	0.58	0.23	19%	100%	0.05	0.76	−0.01	6%	100%	0.01	0.59	0.17	3%	100%	0.03
Five-years-ahead	36	0.62	0.34	56%	67%	0.14	0.70	0.32	47%	86%	0.17	0.67	0.32	44%	64%	0.11

Notes: This table displays the average estimates of α and β based on the models in Eqns. (4.13) and (4.14) as well as the average across the R^2-statistics of all specifications. We also report the share of estimated slope coefficients that are significantly different from (i) zero ($\beta = 0$) and (ii) one ($\beta = 1$) at the 5% critical level against two-sided alternatives. The overall results in the first row are based on all 108 specifications of Eqns. (4.13) and (4.14). The rows for the employed dispersion statistic and the disagreement measure are based on 54 specifications each and the remaining rows are based on 36 specifications each. The results for the main results (columns 3–7) include all observations from 1999Q1 to 2018Q2. The estimates for the sample split (columns 8–12) are based on the period 2008Q2–2018Q2. Columns 13–17 display the results for the full sample period when only the 50% most regular SPF participants are considered in the calculation of the disagreement and uncertainty measures from Section 4.2.

existence of a strong link between uncertainty and disagreement in the SPF data.

As the next step, we analyze the conditional results. In line with the findings of Abel et al. (2016), the average estimates of the slope coefficient are moderately larger for the 54 estimates based on the variance-based measures compared to the 54 estimates using the IQR statistics (0.29 vs. 0.21). This finding implies that studies that rely on variance-based measures may be affected by influential observations in the cross-section. In particular, this may be the case for Zarnowitz and Lambros (1987), who document large correlations between disagreement and uncertainty but use a relatively short time series in their analysis.

The average estimate of β is smaller if disagreement is calculated from the point forecasts rather than the means of the respective distribution (0.19 vs. 0.30). Notably, there is a substantial difference in the share of statistically significant estimates of the slope coefficient in this case (22% vs. 46%). Thus, measures of disagreement derived from information contained in the histograms are more closely related to uncertainty. This may partially explain the relatively large correlations documented in Boero et al. (2015) and Glas and Hartmann (2016), among others. However, as mentioned in Section 4.2, the primary advantage of using disagreement as a proxy variable for uncertainty is that only point forecasts are required in the construction of this statistic. If the researcher has access to the subjective probability distributions, it seems recommendable to directly compute uncertainty instead. Moreover, Clements (2010) shows that the point forecasts are more accurate than the histogram means.

In contrast to the previous two dimensions, distributional assumptions appear to be of little importance. The estimates of β are almost identical regardless of whether the 'mass-at-midpoint' approach, uniformly distributed probabilities or beta distributions are assumed (0.26 vs. 0.24 vs. 0.25). The shares of significant estimates as well as the averages across the goodness of fit are also very similar.

The average slope coefficients also fall within a relatively narrow range if distinct outcome variables are considered. However, we obtain more than twice as many significant estimates of β and a higher goodness of fit in the case of the unemployment rate compared to the other variables. This is noteworthy because predictions of the unemployment rate are less frequently the focus of empirical research compared to the inflation rate or output growth. A notable exception is Abel et al. (2016), who also document a relatively strong link between uncertainty and disagreement for the predictions of the unemployment rate reported in the SPF. The hypothesis that $\beta = 1$ is rejected for all specfications in the case of the expected unemployment rate.

Lastly, we find that the strength of the link between disagreement and uncertainty increases with the forecast horizon. This finding squares with Zarnowitz and Lambros (1987) but is at variance with the evidence from Lahiri and Sheng (2010) and Glas and Hartmann (2016), who show that disagreement and uncertainty deviate more strongly for large forecast horizons. Inspection of Tables 4.6 and 4.7 reveals that most of the significant estimates of β for $h = 20$ are obtained for the predictions of output growth and the unemployment rate, which are not considered in Glas and Hartmann (2016). The analysis of Lahiri and Sheng (2010) is based on specific model assumptions and disregards the predictions of the unemployment rate. As shown in Figures 4.3 and 4.4, disagreement about the expected unemployment rate increases with the forecast horizon. This finding suggests that the strength of the uncertainty-disagreement linkage may depend on the persistence of the considered outcome variable.

In sum, we find little evidence for the existence of a robust link between disagreement and uncertainty in the SPF data. Across all specifications, the average estimates of the intercept and the slope coefficient strongly deviate from the case where $\alpha = 0$ and $\beta = 1$. Moreover, the R^2-statistics are extremely small throughout, suggesting that forecaster disagreement captures very little of the movement in the uncertainty series. Out of the considered dimensions of interest, the dispersion statistic, the considered disagreement measure, the outcome variable and the forecast horizon all have an impact on the strength of the link to some degree. In contrast, distributional assumptions do not appear to be very influential. We proceed by assessing the robustness of our results to the considered sample period below.

4.4.2 Uncertainty During and After the Great Recession

In order to assess the robustness of the main results, we estimate each model on a subsample of observations that covers the years following the outbreak of the Great Recession. The aim is to investigate whether the link between disagreement and uncertainty differs systematically during economically calm and turbulent periods (Boero et al., 2015). Research on the importance of macroeconomic uncertainty has regained prominence with the outbreak of the financial crisis in 2008. The subsequent European sovereign debt crisis was characterized by extremely low interest rates and considerable economic and political instability (Baker et al., 2016). More recently, the outcome of the UK referendum in 2016Q2 and the election of Donald Trump as the president of the United States in 2017Q1 may have had a significant impact on the economic outlook in the Euro area. In light of these developments, we assess the robustness of the main findings during this particular period. Columns 8–12 of Table 4.3 present the estimates of Eqns. (4.13) and (4.14) based on a smaller sample of observations covering the period 2008Q2–2018Q2. We choose 2008Q2 as the beginning of the subsample because it represents the first CEPR-based recession period in our data and roughly coincides with the outbreak of the Great Recession. However, the results are robust to alternative choices regarding the beginning of the subsample.

The average estimates of the intercept and the slope coefficient across all 108 models are equal to 0.71 and 0.10, respectively. Compared to the full-sample estimates, the share of significant slope coefficients is almost halved in this case (34% vs. 19%), whereas the share of rejections of the hypothesis that $\beta = 1$ increases from 89% to 95%. Thus, the relationship between disagreement and uncertainty is considerably weaker during the more turbulent periods in our sample.

Turning to the results of the conditional specifications, we find that the link is moderately stronger for the IQR- than for the variance-based measures, whereas the choice of the considered disagreement statistic is of little importance. These patterns differ from the full-sample results. As before, there is little evidence of substantial differences across the results for the employed distributional assumptions. With respect to the outcome variable, the smallest average estimate of β is obtained for the predictions of the unemployment rate. The share of significant slope coefficients is also considerably smaller in this case. Thus, the difference to the full-sample results is primarily ascribed to a weakened link between disagreement and uncertainty in the case of the expected unemployment rate during economically turbulent periods. With respect to the forecast horizon, it is particularly striking that the average estimates of β for the one-year- and two-years-ahead forecasts are *negative*, albeit all of the individual negative estimates are insignificant (see Tables 4.8–4.9). This finding suggests that disagreement is a particularly bad proxy at short forecast horizons during periods of great economic instability. In contrast, the link remains positive and (relatively) stable for the long-term predictions.

Overall, we document that the link between disagreement and uncertainty is particularly weak during periods when measuring the latter is most relevant. This is in line with Lahiri and Sheng (2010), who find that disagreement is not a valid indicator of uncertainty during turbulent periods. The R^2-statistics remain at relatively low levels in most cases except for the five-years-ahead predictions.

4.4.3 Entry and Exit of Survey Participants

In a second robustness check, we assess the importance of the entry and exit of individual participants to and from the survey panel. The evidence from Figure 4.1 shows that participation in the SPF declines over time, thus providing substantial evidence for attrition. It may be the case that forecasters with a relatively poor forecast performance decide to leave the survey, resulting in a positive selection of the remaining cross-section

(Capistran and Timmermann, 2009). Moreover, a seasonal pattern in SPF participation is documented in Table 4.2. López-Pérez (2016) shows that forecaster participation is related to the level of uncertainty, suggesting that systematic exit from the SPF panel may bias the employed measures. For this reason, Lahiri and Liu (2006) and Abel et al. (2016) propose to focus on a smaller set of regular survey participants. Hence, we present the estimates of Eqns. (4.13) and (4.14) when only SPF participants with sufficiently high participation rates are considered in the calculation of the measures from Section 4.2.

Regular participants are classified in the following way: For each combination of outcome variable and forecast horizon, we calculate the individual participation rate of each forecaster as the number of times forecaster i submitted matched point and histogram forecasts divided by the total number of survey rounds.[12] Next, we compute the median across the individual participation rates and calculate the disagreement and uncertainty statistics from Section 4.2 using only the predictions of panelists with individual participation rates that are at least as large as the respective median rate.[13] Thus, for each outcome variable and forecast horizon, only the 50% most regular SPF participants are considered. This is considerably more restrictive than the studies of D'Amico and Orphanides (2008) and Capistran and Timmermann (2009), who both consider minimum participation rates of approximately 10%. Based on the re-calculated measures, we estimate Eqns. (4.13) and (4.14).

The results in columns 13–17 of Table 4.3 are relatively similar to the findings from the full sample of observations in columns 3–7. The average estimates of α and β based on all 108 estimates are 0.60 and 0.21, respectively. The slope coefficient is thus slightly smaller than the one obtained for the full cross-section. The share of significant estimates of β as well as the average R^2 are also smaller than before.

The conditional results for the employed dispersion statistic, disagreement measure and distribution are broadly comparable to the full-sample results, albeit smaller in terms of magnitude. Interestingly, the share of significant slope coefficients is almost identical across outcome variables, whereas for the full sample more than twice as many significant estimates are obtained in the case of the unemployment rate compared to the other variables. It may the case that the previously documented strong link between disagreement and uncertainty in the case of the expeted unemployment rate is the result of extreme predictions reported by a few irregular SPF participants. With respect to the forecast

[12] Our sample includes 78 surveys for the one-year- and two-years-ahead forecasts and 72 surveys for the long-term projections. Additionally, we exclude the 2009Q1 survey for the real GDP growth forecasts.

[13] Median participation rates for inflation are 67%, 64% and 67% for the one-year, two-years- and five-years-ahead forecasts, respectively. The corresponding rates for the other variables are 64%, 64% and 67% for output growth as well as 64%, 63% and 67% for the unemployment rate, respectively.

horizon, it is particularly the two-years-ahead forecasts that tend to lose their significance. The relationship for the long-term predictions remains relatively strong.

To summarize, the results from this robustness exercise suggest that the weak evidence regarding the existence of a robust relationship between disagreement and uncertainty is not just a result of a bias introduced by systematic entry and exit of forecasters to and from the SPF panel. In fact, most of the results are slightly *weaker* than in the case of the full sample, which suggests that the relatively modest evidence that is documented in the empirical literature may still be overstated to some extent.

4.5 Comparison with Other Measures of Uncertainty

In Section 4.4 we have focused on the relationship between survey-based indicators of uncertainty. However, various other proxy variables have been proposed in the literature. In this section, we analyze how closely the survey-based measures are related to these alternative quantifications. This will help us to understand the divergence between uncertainty and disagreement documented in Figure 4.7.

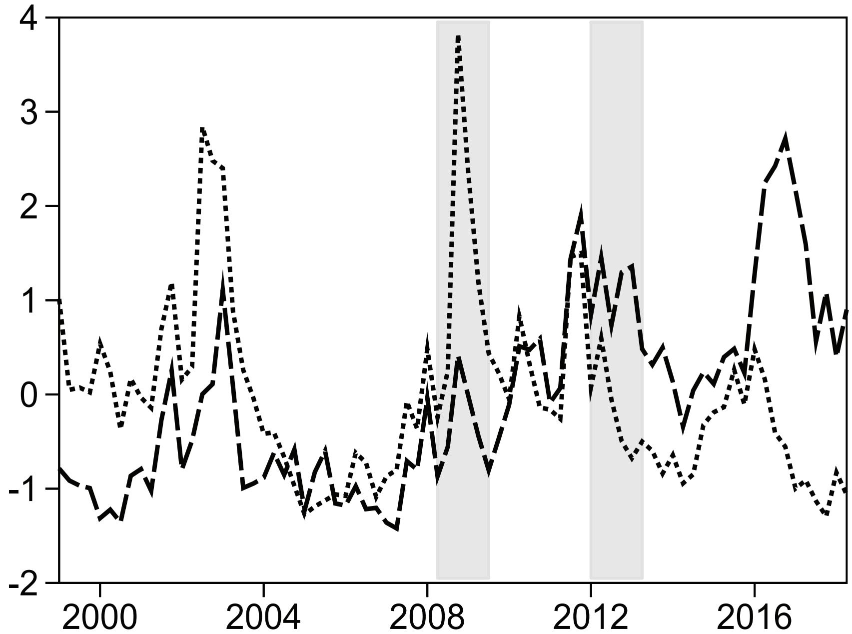
One of the most popular indicators is the Economic Policy Uncertainty (EPU) index by Baker et al. (2016), which captures newspaper coverage of topics related to macroeconomic uncertainty. The EPU is a standardized measure that counts the number of newspaper articles containing certain key terms related to economic uncertainty and captures the long-term outlook of economic agents. We use the quarterly average over the monthly European index.

Another common approach is to consider the fluctuations on financial markets. Schwert (1989) documents a positive, but weak, association between the volatilities of the stock market and macroeconomic variables. Realized stock market volatility is a backward-looking measure since it is based on past returns that are observed ex-post. In order to analyze market participants' expectations regarding the movements on stock markets, we use the VSTOXX volatility index, which captures the implied volatility on the European stock market over the next 30 days. This index is based on the EuroStoxx 50 options and conceptually similar to the CBOE Volatility Index (VXO). In contrast to realized volatility, the VSTOXX is forward-looking and captures the uncertainty on financial markets. Therefore, it is more comparable to the ex-ante measures of uncertainty derived from the SPF data. We use the quarterly average over the daily observations.

In the general equilibrium model of Pástor and Veronesi (2013) stock prices respond to political news, which establishes a theoretical link between financial and policy uncertainty. One implication of their model is that the EPU and VSTOXX indices should be positively related. A link between the two measures can also be motivated based on the fact that they are both associated with the overall economic outlook. For example, market participants may postpone investments in financial assets during periods characterized by high policy uncertainty, which should be associated with an increase in expected stock volatility. Empirically, Henzel and Rengel (2017) conduct a factor analysis to show that both the EPU and the VXO are closely related to uncertainty about the business cycle.

To examine the comovement between policy and financial uncertainty, Figure 4.8 depicts the time series for EPU and VSTOXX over the period 1999Q1–2018Q2. We standardize both series to have mean zero and variance one.

Figure 4.8: Policy versus financial uncertainty



Notes: This plot depicts the standardized quarterly time series of the Economic Policy Uncertainty (EPU, dashed line) and VSTOXX indices (dotted line). The sample covers the period 1999Q1–2018Q2. Shaded gray areas indicate CEPR-based recession periods.

Before the outbreak of the sovereign debt crisis, the EPU and VSTOXX indices comove, although the VSTOXX is more volatile than the EPU, e.g., during the outbreak of the Great Recession. A decoupling of both measures is visible since around 2012. While EPU has risen to relatively high levels, the uncertainty on financial markets remains rather low. It could be the case that major political events such as the outcome of the Brexit referendum and the election of Donald Trump may be responsible for the increase in the EPU. At the same time, benign global economic conditions and the impact of loose monetary policy may have caused the VSTOXX to remain at rather low levels.

The observed divergence of EPU and VSTOXX coincides with the deviation of forecaster disagreement and survey-based uncertainty documented in Figure 4.7. It may be the case that the same factors that are responsible for the divergence of EPU and VSTOXX in recent years are responsible for the large deviations of the measures from the SPF.

To analyze whether the survey-based indicators are more closely related to long-term policy uncertainty or short-term financial uncertainty, we calculate the bivariate correlation statistics of each measure of disagreement and uncertainty with the quarterly (non-standardized) averages of either the EPU or the VSTOXX index. The results are summarized in Table 4.4.[14]

The evidence suggests that the six survey-based measures of uncertainty are strongly related to overall policy uncertainty, as measured by the EPU index. The magnitude of the correlations is relatively similar regardless of the considerd statistic, distributional assumption, outcome variable and forecast horizon. In contrast, the correlation statistics between the disagreement measures and EPU are rather small. In particular, we can only partially confirm the positive association between forecaster disagreement and newspaper-based uncertainty that has been documented by Dovern (2015). In some instances, the correlation statistics are even negative. Thus, we conclude that survey-based uncertainty is closely related to overall policy uncertainty, whereas forecaster disagreement is not.

A different picture emerges for the relationship between the survey-based measures and financial uncertainty. The correlation statistics between the six uncertainty measures and the VSTOXX index are small and negative in most cases. In contrast, the association between VSTOXX and disagreement is positive and strong at the short and medium horizons. This finding suggests that forecaster disagreement is closely related to the expected fluctuations on financial markets. For the long-term predictions, the correlation statistics are negative in many cases. Based on factor analysis, Sauter (2012) finds a similarly strong connection between short-term forecaster disagreement based on the SPF data and the VSTOXX index. Glas and Hartmann (2016) document that the strength of the relationship between disagreement and realized volatility declines with the forecast horizon. We consider the forward-looking VSTOXX index instead. Nonetheless, we also

[14]The deadline for the submission of the SPF questionnaire is usually around the middle or the end of the first month of each quarter. Thus, the observations from the remaining months (EPU) or days (VSTOXX) are unknown to the SPF participants when they report their predictions. To analyze whether this inaccurate representation of the forecasters' information set affects the results of the correlation analysis, we have considered alternative definitions of EPU and VSTOXX besides taking the quarterly average. Instead, we calculated EPU and VSTOXX in quarter t as the average across all observations between the deadline for the previous survey round from quarter $t-1$ and the deadline for the current round (excluding the former but including the latter). The bivariate correlations of the alternative EPU/VSTOXX indices with the survey-based measures were almost identical to the ones reported in Table 4.4.

Table 4.4: Correlations of survey-based and other uncertainty proxies

		Inflation rate			GDP growth rate			Unemployment rate		
Horizon h		4	8	20	4	8	20	4	8	20
EPU	$\overline{\sigma}^{\mathrm{M}}_{t+h\|t}$	0.79	0.79	0.73	0.73	0.75	0.72	0.76	0.76	0.69
	$\overline{\sigma}^{\mathrm{U}}_{t+h\|t}$	0.79	0.79	0.73	0.73	0.75	0.72	0.76	0.76	0.69
	$\overline{\sigma}^{\mathrm{B}}_{t+h\|t}$	0.79	0.79	0.72	0.74	0.74	0.72	0.76	0.76	0.69
	$\tilde{\phi}^{\mathrm{M}}_{t+h\|t}$	0.73	0.76	0.68	0.70	0.67	0.69	0.74	0.69	0.61
	$\tilde{\phi}^{\mathrm{U}}_{t+h\|t}$	0.78	0.76	0.66	0.62	0.65	0.60	0.71	0.69	0.64
	$\tilde{\phi}^{\mathrm{B}}_{t+h\|t}$	0.77	0.75	0.64	0.61	0.65	0.60	0.71	0.72	0.64
	$s_{t+h\|t}$	0.22	0.09	0.07	−0.05	−0.16	0.20	0.13	0.14	0.19
	$s^{\star}_{t+h\|t}$	0.23	0.09	0.37	−0.01	−0.06	0.27	0.14	0.23	0.22
	$s^{\mathrm{B}}_{t+h\|t}$	0.23	0.08	0.35	−0.01	−0.07	0.28	0.13	0.23	0.22
	$\tilde{f}_{t+h\|t}$	0.02	0.08	−0.11	−0.07	−0.08	0.15	−0.02	0.07	0.14
	$\tilde{f}^{\star}_{t+h\|t}$	0.07	0.11	0.51	0.02	−0.01	−0.05	0.08	0.17	0.21
	$\tilde{f}^{\mathrm{B}}_{t+h\|t}$	0.08	0.09	0.51	0.01	−0.02	−0.05	0.08	0.16	0.22
VSTOXX	$\overline{\sigma}^{\mathrm{M}}_{t+h\|t}$	−0.07	−0.05	−0.18	−0.06	0.03	−0.11	−0.20	−0.16	−0.30
	$\overline{\sigma}^{\mathrm{U}}_{t+h\|t}$	−0.07	−0.05	−0.18	−0.06	0.03	−0.11	−0.20	−0.16	−0.30
	$\overline{\sigma}^{\mathrm{B}}_{t+h\|t}$	−0.07	−0.05	−0.17	−0.04	0.06	−0.08	−0.19	−0.14	−0.29
	$\tilde{\phi}^{\mathrm{M}}_{t+h\|t}$	−0.03	0.05	−0.18	−0.04	0.06	−0.02	−0.11	−0.11	−0.27
	$\tilde{\phi}^{\mathrm{U}}_{t+h\|t}$	−0.03	−0.03	−0.13	0.05	0.03	−0.17	−0.11	−0.12	−0.18
	$\tilde{\phi}^{\mathrm{B}}_{t+h\|t}$	0.00	−0.01	−0.15	0.04	0.04	−0.15	−0.15	−0.14	−0.22
	$s_{t+h\|t}$	0.52	0.25	0.15	0.44	0.40	−0.39	0.45	0.38	0.11
	$s^{\star}_{t+h\|t}$	0.52	0.28	0.13	0.45	0.39	−0.13	0.38	0.30	0.11
	$s^{\mathrm{B}}_{t+h\|t}$	0.52	0.29	0.14	0.46	0.40	−0.15	0.38	0.30	0.11
	$\tilde{f}_{t+h\|t}$	0.44	0.23	0.02	0.42	0.29	−0.49	0.45	0.36	−0.18
	$\tilde{f}^{\star}_{t+h\|t}$	0.38	0.29	−0.10	0.49	0.46	−0.18	0.33	0.23	0.11
	$\tilde{f}^{\mathrm{B}}_{t+h\|t}$	0.38	0.31	−0.11	0.48	0.47	−0.22	0.34	0.22	−0.12

Notes: For each outcome variable and forecast horizon, this table displays the bivariate correlation between one particular measure of forecaster disagreement or survey-based uncertainty and the quarterly average of either the Economic Policy Uncertainty (top panel) or the VSTOXX volatility indices (bottom panel). The sample period is 1999Q1–2018Q2.

find that the magnitude of the association diminishes in the forecast horizon.

Overall, the results of the correlation analysis support the notion that uncertainty and disagreement are fundamentally different concepts. Survey-based uncertainty is associated with overall policy uncertainty, whereas forecaster disagreement is more closely related to the volatility on financial markets. The same factors that are responsible for the divergence of policy and financial uncertainty may thus also play a role in the weakened relationship between forecaster disagreement and survey-based uncertainty in recent years.

4.6 Conclusion

We analyze the relationship between macroeconomic uncertainty and forecaster disagreement in the Euro area using data from the European Central Bank's Survey of Professional Forecasters for the period 1999Q1–2018Q2. In addition to the overall strength of this link, we investigate the importance of the employed dispersion statistic, the choice between point forecasts and histogram means to calculate disagreement, commonly used distributional assumptions, the considered outcome variable and the forecast horizon. In line with previous studies, we find that disagreement is not a good proxy for uncertainty. However, the conditional results allow us to shed additional light on the mechanics behind this relationship. Among the stylized facts described in this chapter, we find that both the employed dispersion statistic and the disagreement measure are relevant factors in the sense that the link is stronger for variance-based statistics and disagreement based on the histogram means. The considered outcome variable and forecast horizon are also of some relevance. In contrast, distributional assumptions appear to have little influence. Our results suggest that the results from empirical research in this field should be interpreted within the context of the specific choices and assumptions by the researcher.

We also account for differences in the results during economically turbulent periods and the impact of the entry and exit of forecasters to and from the SPF panel. The results show that the link between disagreement and uncertainty is particularly weak during economically turbulent periods, i.e., at times when such measures are most relevant. In contrast, the entry and exit of survey participants has little impact on our findings. Moreover, we show that survey-based measures of uncertainty are associated with overall policy uncertainty, whereas disagreement is more closely related to the expected fluctuations on financial markets. Thus, the two concepts are fundamentally different. The same factors that explain the divergence between policy and financial uncertainty may be responsible for the particularly weak relationship between forecaster disagreement and survey-based uncertainty in recent years.

4.7 Appendix

Table 4.5: Uncertainty and disagreement (replication of Abel et al., 2016)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
	Panel A: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$								
α	0.39*	0.48*	0.58*	0.45*	0.46*	0.55*	0.40*	0.40*	0.58*
	(0.03)	(0.04)	(0.06)	(0.03)	(0.06)	(0.04)	(0.03)	(0.03)	(0.08)
β	0.46*	0.38*	0.26	0.26*	0.54*	0.54*	0.33*	0.38*	0.20
	(0.10)	(0.18)	(0.31)	(0.05)	(0.18)	(0.19)	(0.09)	(0.10)	(0.16)
R^2	0.20	0.13	0.06	0.22	0.15	0.14	0.19	0.27	0.12
	Panel B: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$								
α	0.61*	0.64*	0.83*	0.56*	0.59*	0.74*	0.52*	0.56*	0.71*
	(0.06)	(0.07)	(0.08)	(0.06)	(0.05)	(0.04)	(0.05)	(0.05)	(0.07)
β	0.07	0.18	−0.41	0.23*	0.40*	0.28	0.21*	0.21*	0.15
	(0.14)	(0.28)	(0.30)	(0.08)	(0.08)	(0.19)	(0.05)	(0.07)	(0.11)
R^2	0.01	0.03	0.15	0.09	0.12	0.07	0.10	0.10	0.07

Notes: This table displays the estimates from a replication of Table II from Abel et al. (2016). The sample period $t = 1, \ldots, 60$ represents time instances between 1999Q1 and 2013Q4. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.6: Uncertainty and disagreement (full-sample estimates, point forecasts)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
	Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$								
α	0.42*	0.51*	0.61*	0.54*	0.63*	0.49*	0.47*	0.48*	0.61*
	(0.06)	(0.07)	(0.09)	(0.08)	(0.11)	(0.04)	(0.07)	(0.08)	(0.10)
β	0.36*	0.31	0.14	0.08	0.03	0.78*	0.15	0.25	0.18
	(0.13)	(0.25)	(0.40)	(0.15)	(0.31)	(0.19)	(0.15)	(0.13)	(0.16)
R^2	0.10	0.06	0.01	0.01	0.00	0.24	0.02	0.07	0.07
	Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$								
α	0.45*	0.53*	0.63*	0.55*	0.65*	0.51*	0.49*	0.50*	0.63*
	(0.06)	(0.07)	(0.08)	(0.07)	(0.11)	(0.04)	(0.07)	(0.08)	(0.10)
β	0.35*	0.30	0.14	0.08	0.03	0.76*	0.14	0.24	0.18
	(0.12)	(0.24)	(0.39)	(0.14)	(0.30)	(0.18)	(0.15)	(0.12)	(0.16)
R^2	0.10	0.06	0.01	0.01	0.00	0.24	0.02	0.07	0.08
	Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$								
α	0.39*	0.47*	0.56*	0.49*	0.58*	0.45*	0.43*	0.44*	0.56*
	(0.06)	(0.07)	(0.08)	(0.07)	(0.11)	(0.04)	(0.07)	(0.08)	(0.09)
β	0.34*	0.30	0.16	0.09	0.05	0.75*	0.16	0.26*	0.19
	(0.13)	(0.24)	(0.39)	(0.14)	(0.30)	(0.18)	(0.14)	(0.12)	(0.14)
R^2	0.09	0.06	0.02	0.02	0.00	0.24	0.02	0.09	0.10
	Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$								
α	0.58*	0.63*	0.77*	0.64*	0.73*	0.76*	0.57*	0.59*	0.67*
	(0.05)	(0.08)	(0.09)	(0.10)	(0.08)	(0.03)	(0.08)	(0.08)	(0.04)
β	0.11	0.25	−0.08	0.06	0.10	0.26	0.10	0.22	0.24*
	(0.13)	(0.29)	(0.43)	(0.16)	(0.18)	(0.14)	(0.13)	(0.11)	(0.08)
R^2	0.01	0.04	0.00	0.01	0.01	0.07	0.01	0.06	0.17
	Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$								
α	0.64*	0.69*	0.81*	0.63*	0.72*	0.73*	0.59*	0.62*	0.70*
	(0.06)	(0.09)	(0.09)	(0.06)	(0.08)	(0.03)	(0.07)	(0.08)	(0.05)
β	0.06	0.13	−0.15	0.10	0.13	0.40*	0.10	0.18	0.21*
	(0.13)	(0.29)	(0.42)	(0.10)	(0.17)	(0.15)	(0.10)	(0.10)	(0.09)
R^2	0.00	0.01	0.01	0.02	0.02	0.13	0.02	0.05	0.12
	Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$								
α	0.60*	0.65*	0.82*	0.60*	0.70*	0.72*	0.54*	0.57*	0.68*
	(0.07)	(0.11)	(0.11)	(0.08)	(0.10)	(0.03)	(0.08)	(0.10)	(0.05)
β	0.07	0.17	−0.21	0.11	0.16	0.42*	0.10	0.22	0.25*
	(0.17)	(0.35)	(0.50)	(0.13)	(0.22)	(0.17)	(0.13)	(0.12)	(0.08)
R^2	0.00	0.01	0.02	0.02	0.02	0.11	0.01	0.05	0.14

Notes: This table displays the full-sample estimates of Eqns. (4.13) and (4.14) using disagreement statistics based on the point forecasts for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \dots, 78$ represents time instances between 1999Q1 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.7: Uncertainty and disagreement (full-sample estimates, histogram means)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$									
α	0.41*	0.49*	0.52*	0.52*	0.59*	0.53*	0.46*	0.44*	0.56*
	(0.06)	(0.07)	(0.07)	(0.07)	(0.10)	(0.07)	(0.08)	(0.07)	(0.07)
β	0.41*	0.37	0.52	0.12	0.16	0.62*	0.17	0.35*	0.28*
	(0.12)	(0.26)	(0.34)	(0.12)	(0.28)	(0.22)	(0.16)	(0.09)	(0.09)
R^2	0.12	0.08	0.15	0.03	0.01	0.18	0.02	0.13	0.17
Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$									
α	0.44*	0.51*	0.53*	0.54*	0.61*	0.55*	0.49*	0.46*	0.57*
	(0.05)	(0.07)	(0.07)	(0.07)	(0.10)	(0.07)	(0.07)	(0.06)	(0.07)
β	0.39*	0.36	0.51	0.11	0.15	0.61*	0.16	0.34*	0.27*
	(0.12)	(0.25)	(0.33)	(0.12)	(0.28)	(0.22)	(0.16)	(0.09)	(0.09)
R^2	0.12	0.08	0.15	0.03	0.01	0.18	0.02	0.13	0.17
Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s^{\mathrm{B}}_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$									
α	0.38*	0.46*	0.48*	0.48*	0.55*	0.48*	0.42*	0.39*	0.51*
	(0.06)	(0.07)	(0.07)	(0.07)	(0.10)	(0.07)	(0.07)	(0.06)	(0.07)
β	0.38*	0.34	0.50	0.13	0.16	0.62*	0.18	0.36*	0.27*
	(0.12)	(0.26)	(0.34)	(0.12)	(0.28)	(0.20)	(0.16)	(0.09)	(0.09)
R^2	0.12	0.07	0.15	0.04	0.01	0.18	0.03	0.14	0.19
Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$									
α	0.54*	0.59*	0.55*	0.61*	0.69*	0.85*	0.52*	0.53*	0.54*
	(0.06)	(0.07)	(0.07)	(0.10)	(0.08)	(0.06)	(0.07)	(0.06)	(0.07)
β	0.24*	0.38	0.71*	0.16	0.23	−0.02	0.23*	0.32*	0.41*
	(0.09)	(0.23)	(0.24)	(0.16)	(0.17)	(0.19)	(0.09)	(0.07)	(0.07)
R^2	0.05	0.09	0.23	0.04	0.03	0.00	0.07	0.13	0.38
Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$									
α	0.61*	0.64*	0.61*	0.59*	0.69*	0.85*	0.55*	0.58*	0.55*
	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)
β	0.15	0.27	0.61*	0.20	0.21	0.01	0.19*	0.24*	0.40*
	(0.10)	(0.22)	(0.22)	(0.11)	(0.15)	(0.20)	(0.08)	(0.08)	(0.06)
R^2	0.03	0.05	0.21	0.08	0.03	0.00	0.07	0.10	0.35
Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\mathrm{B}}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$									
α	0.55*	0.60*	0.59*	0.55*	0.68*	0.85*	0.49*	0.53*	0.52*
	(0.07)	(0.09)	(0.09)	(0.08)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)
β	0.21	0.32	0.67*	0.24	0.20	0.01	0.21*	0.30*	0.44*
	(0.13)	(0.28)	(0.33)	(0.13)	(0.17)	(0.24)	(0.09)	(0.08)	(0.05)
R^2	0.03	0.04	0.16	0.08	0.02	0.00	0.06	0.10	0.36

Notes: This table displays the full-sample estimates of Eqns. (4.13) and (4.14) using disagreement statistics based on the histogram means for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \ldots, 78$ represents time instances between 1999Q1 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.8: Uncertainty and disagreement (sample split, point forecasts)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
	Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$								
α	0.62*	0.65*	0.64*	0.65*	0.75*	0.58*	0.64*	0.74*	0.69*
	(0.02)	(0.04)	(0.06)	(0.02)	(0.02)	(0.05)	(0.03)	(0.04)	(0.10)
β	−0.05	0.05	0.28	−0.01	−0.08	0.58*	−0.11	−0.09	0.16
	(0.08)	(0.14)	(0.23)	(0.04)	(0.09)	(0.15)	(0.07)	(0.07)	(0.11)
R^2	0.01	0.01	0.10	0.00	0.02	0.26	0.06	0.04	0.13
	Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$								
α	0.64*	0.67*	0.65*	0.66*	0.76*	0.60*	0.65*	0.75*	0.71*
	(0.02)	(0.04)	(0.06)	(0.02)	(0.02)	(0.05)	(0.02)	(0.04)	(0.10)
β	−0.04	0.05	0.27	−0.01	−0.08	0.57*	−0.11	−0.09	0.16
	(0.07)	(0.13)	(0.22)	(0.04)	(0.09)	(0.14)	(0.07)	(0.07)	(0.10)
R^2	0.01	0.01	0.10	0.00	0.02	0.26	0.06	0.04	0.13
	Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$								
α	0.58*	0.60*	0.59*	0.60*	0.70*	0.54*	0.59*	0.68*	0.64*
	(0.02)	(0.04)	(0.06)	(0.02)	(0.02)	(0.05)	(0.02)	(0.04)	(0.09)
β	−0.05	0.05	0.27	−0.00	−0.07	0.58*	−0.10	−0.07	0.17
	(0.08)	(0.12)	(0.22)	(0.03)	(0.09)	(0.15)	(0.06)	(0.07)	(0.10)
R^2	0.01	0.01	0.11	0.00	0.02	0.27	0.05	0.02	0.14
	Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$								
α	0.71*	0.75*	0.74*	0.77*	0.87*	0.79*	0.74*	0.84*	0.76*
	(0.03)	(0.03)	(0.05)	(0.02)	(0.02)	(0.07)	(0.02)	(0.04)	(0.06)
β	0.05	0.19*	0.45*	0.02	−0.01	0.35	−0.05	−0.02	0.22*
	(0.09)	(0.06)	(0.19)	(0.03)	(0.09)	(0.18)	(0.04)	(0.05)	(0.04)
R^2	0.01	0.11	0.18	0.00	0.00	0.19	0.01	0.00	0.31
	Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$								
α	0.76*	0.83*	0.79*	0.72*	0.84*	0.73*	0.73*	0.85*	0.78*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)	(0.08)	(0.01)	(0.04)	(0.05)
β	−0.03	0.00	0.29	0.09	0.06	0.56*	−0.02	−0.06	0.20*
	(0.06)	(0.05)	(0.18)	(0.06)	(0.08)	(0.22)	(0.03)	(0.06)	(0.05)
R^2	0.00	0.00	0.09	0.06	0.01	0.28	0.00	0.02	0.25
	Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$								
α	0.74*	0.83*	0.80*	0.71*	0.85*	0.71*	0.70*	0.85*	0.79*
	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)	(0.09)	(0.02)	(0.04)	(0.06)
β	−0.02	0.00	0.28	0.09	0.05	0.65*	−0.05	−0.06	0.22*
	(0.08)	(0.06)	(0.21)	(0.08)	(0.09)	(0.24)	(0.04)	(0.06)	(0.04)
R^2	0.00	0.00	0.07	0.03	0.01	0.29	0.01	0.02	0.25

Notes: This table displays the reduced-sample estimates of Eqns. (4.13) and (4.14) using disagreement statistics based on the point forecasts for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \ldots, 41$ represents time instances between 2008Q2 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.9: Uncertainty and disagreement (sample split, histogram means)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$									
α	0.62*	0.66*	0.57*	0.64*	0.74*	0.67*	0.64*	0.72*	0.72*
	(0.02)	(0.04)	(0.09)	(0.02)	(0.03)	(0.04)	(0.03)	(0.05)	(0.09)
β	−0.05	0.01	0.47	0.00	−0.07	0.27*	−0.12	−0.04	0.13
	(0.08)	(0.12)	(0.34)	(0.04)	(0.12)	(0.11)	(0.07)	(0.09)	(0.10)
R^2	0.01	0.00	0.17	0.00	0.01	0.11	0.05	0.01	0.08
Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$									
α	0.64*	0.68*	0.59*	0.66*	0.76*	0.69*	0.66*	0.73*	0.74*
	(0.02)	(0.04)	(0.09)	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.09)
β	−0.05	0.01	0.46	0.00	−0.07	0.27*	−0.11	−0.04	0.13
	(0.08)	(0.11)	(0.33)	(0.03)	(0.12)	(0.11)	(0.07)	(0.09)	(0.09)
R^2	0.01	0.00	0.17	0.00	0.01	0.11	0.06	0.01	0.08
Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s^{\mathrm{B}}_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$									
α	0.58*	0.62*	0.54*	0.60*	0.70*	0.62*	0.59*	0.66*	0.66*
	(0.02)	(0.04)	(0.08)	(0.02)	(0.02)	(0.04)	(0.03)	(0.04)	(0.08)
β	−0.06	0.01	0.43	0.01	−0.06	0.27*	−0.10	−0.02	0.14
	(0.08)	(0.10)	(0.32)	(0.03)	(0.11)	(0.12)	(0.07)	(0.07)	(0.09)
R^2	0.01	0.00	0.16	0.00	0.01	0.11	0.04	0.00	0.10
Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$									
α	0.73*	0.76*	0.67*	0.77*	0.87*	0.88*	0.72*	0.78*	0.73*
	(0.03)	(0.03)	(0.07)	(0.02)	(0.04)	(0.04)	(0.03)	(0.05)	(0.07)
β	−0.01	0.13*	0.51*	0.03	−0.01	0.10	0.01	0.08	0.27*
	(0.09)	(0.06)	(0.20)	(0.04)	(0.14)	(0.12)	(0.04)	(0.08)	(0.06)
R^2	0.00	0.04	0.25	0.01	0.00	0.01	0.00	0.03	0.37
Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$									
α	0.78*	0.85*	0.71*	0.71*	0.84*	0.86*	0.72*	0.82*	0.75*
	(0.03)	(0.03)	(0.06)	(0.03)	(0.05)	(0.06)	(0.02)	(0.05)	(0.08)
β	−0.09	−0.06	0.44*	0.13*	0.08	0.20	−0.01	−0.00	0.25*
	(0.07)	(0.08)	(0.17)	(0.05)	(0.14)	(0.15)	(0.03)	(0.07)	(0.07)
R^2	0.04	0.01	0.24	0.12	0.01	0.03	0.00	0.00	0.31
Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\mathrm{B}}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$									
α	0.76*	0.84*	0.71*	0.69*	0.85*	0.86*	0.69*	0.82*	0.75*
	(0.05)	(0.04)	(0.09)	(0.05)	(0.05)	(0.07)	(0.02)	(0.05)	(0.09)
β	−0.08	−0.03	0.47	0.15	0.05	0.25	−0.01	−0.00	0.27*
	(0.10)	(0.11)	(0.25)	(0.07)	(0.14)	(0.18)	(0.03)	(0.07)	(0.07)
R^2	0.02	0.00	0.18	0.09	0.00	0.03	0.00	0.00	0.30

Notes: This table displays the reduced-sample estimates of Eqns. (4.13) and (4.14) using disagreement statistics based on the histogram means for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \ldots, 41$ represents time instances between 2008Q2 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.10: Uncertainty and disagreement (entry and exit, point forecasts)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
	Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$								
α	0.45*	0.48*	0.72*	0.53*	0.60*	0.51*	0.50*	0.54*	0.69*
	(0.06)	(0.05)	(0.10)	(0.07)	(0.08)	(0.07)	(0.10)	(0.12)	(0.11)
β	0.29*	0.33	−0.29	0.03	0.07	0.79*	0.05	0.11	0.13
	(0.13)	(0.17)	(0.44)	(0.12)	(0.23)	(0.28)	(0.19)	(0.16)	(0.19)
R^2	0.06	0.06	0.04	0.00	0.00	0.29	0.00	0.02	0.03
	Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$								
α	0.47*	0.50*	0.73*	0.55*	0.62*	0.53*	0.52*	0.56*	0.70*
	(0.06)	(0.05)	(0.10)	(0.06)	(0.08)	(0.07)	(0.09)	(0.11)	(0.11)
β	0.28*	0.32	−0.28	0.03	0.07	0.77*	0.05	0.11	0.13
	(0.13)	(0.16)	(0.43)	(0.12)	(0.22)	(0.28)	(0.18)	(0.15)	(0.18)
R^2	0.06	0.06	0.04	0.00	0.00	0.29	0.00	0.02	0.03
	Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$								
α	0.42*	0.45*	0.67*	0.48*	0.55*	0.48*	0.46*	0.49*	0.63*
	(0.06)	(0.05)	(0.10)	(0.06)	(0.08)	(0.07)	(0.09)	(0.11)	(0.10)
β	0.27*	0.31	−0.27	0.04	0.08	0.75*	0.05	0.12	0.13
	(0.12)	(0.16)	(0.42)	(0.12)	(0.22)	(0.25)	(0.18)	(0.15)	(0.17)
R^2	0.05	0.06	0.04	0.01	0.01	0.29	0.00	0.02	0.04
	Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$								
α	0.59*	0.62*	0.82*	0.62*	0.66*	0.72*	0.59*	0.62*	0.85*
	(0.06)	(0.07)	(0.10)	(0.07)	(0.07)	(0.05)	(0.10)	(0.13)	(0.12)
β	0.12	0.20	−0.17	0.04	0.26	0.46*	0.05	0.15	0.07
	(0.07)	(0.20)	(0.39)	(0.12)	(0.17)	(0.15)	(0.13)	(0.12)	(0.14)
R^2	0.02	0.03	0.01	0.00	0.06	0.21	0.00	0.04	0.01
	Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$								
α	0.65*	0.68*	0.84*	0.66*	0.69*	0.69*	0.61*	0.66*	0.81*
	(0.05)	(0.07)	(0.08)	(0.05)	(0.08)	(0.06)	(0.07)	(0.09)	(0.13)
β	0.05	0.09	−0.13	0.01	0.25	0.60*	0.07	0.10	0.13
	(0.06)	(0.19)	(0.36)	(0.10)	(0.17)	(0.18)	(0.09)	(0.08)	(0.14)
R^2	0.00	0.01	0.01	0.00	0.05	0.25	0.01	0.02	0.03
	Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$								
α	0.61*	0.64*	0.84*	0.62*	0.66*	0.67*	0.55*	0.62*	0.83*
	(0.07)	(0.09)	(0.09)	(0.07)	(0.09)	(0.07)	(0.09)	(0.12)	(0.15)
β	0.07	0.17	−0.18	0.04	0.30	0.71*	0.09	0.11	0.13
	(0.07)	(0.21)	(0.39)	(0.12)	(0.19)	(0.20)	(0.12)	(0.11)	(0.16)
R^2	0.00	0.01	0.01	0.00	0.05	0.25	0.01	0.02	0.03

Notes: This table displays the estimates of Eqns. (4.13) and (4.14) using only regular SPF participants and disagreement statistics based on the point forecasts for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \ldots, 78$ represents time instances between 1999Q1 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.11: Uncertainty and disagreement (entry and exit, histogram means)

	Inflation rate			GDP growth rate			Unemployment rate		
Horizon h	4	8	20	4	8	20	4	8	20
	Panel A: $\overline{\sigma}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{M}}_{t+h\|t}$								
α	0.44*	0.49*	0.56*	0.52*	0.61*	0.53*	0.49*	0.50*	0.61*
	(0.06)	(0.05)	(0.07)	(0.06)	(0.08)	(0.09)	(0.10)	(0.10)	(0.08)
β	0.34*	0.29	0.44	0.07	0.05	0.70*	0.10	0.19	0.25*
	(0.12)	(0.18)	(0.31)	(0.11)	(0.25)	(0.31)	(0.19)	(0.12)	(0.11)
R^2	0.09	0.05	0.07	0.02	0.00	0.17	0.01	0.04	0.11
	Panel B: $\overline{\sigma}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times s^{\star}_{t+h\|t} + \varepsilon^{\mathrm{U}}_{t+h\|t}$								
α	0.46*	0.51*	0.58*	0.54*	0.62*	0.55*	0.51*	0.52*	0.63*
	(0.05)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)	(0.10)	(0.10)	(0.08)
β	0.32*	0.28	0.43	0.07	0.05	0.69*	0.09	0.19	0.25*
	(0.12)	(0.17)	(0.30)	(0.10)	(0.24)	(0.31)	(0.18)	(0.12)	(0.11)
R^2	0.09	0.05	0.07	0.02	0.00	0.17	0.01	0.04	0.11
	Panel C: $\overline{\sigma}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times s^{\mathrm{B}}_{t+h\|t} + \varepsilon^{\mathrm{B}}_{t+h\|t}$								
α	0.41*	0.46*	0.53*	0.47*	0.56*	0.49*	0.45*	0.46*	0.57*
	(0.05)	(0.05)	(0.07)	(0.06)	(0.08)	(0.08)	(0.10)	(0.09)	(0.07)
β	0.31*	0.25	0.38	0.08	0.05	0.66*	0.09	0.20	0.24*
	(0.11)	(0.18)	(0.31)	(0.10)	(0.22)	(0.27)	(0.18)	(0.11)	(0.10)
R^2	0.08	0.04	0.06	0.02	0.00	0.19	0.01	0.05	0.12
	Panel D: $\tilde{\phi}^{\mathrm{M}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{M}}_{t+h\|t}$								
α	0.56*	0.60*	0.56*	0.60*	0.71*	0.77*	0.56*	0.60*	0.72*
	(0.05)	(0.06)	(0.09)	(0.08)	(0.06)	(0.08)	(0.09)	(0.10)	(0.09)
β	0.22*	0.27	0.74*	0.09	0.11	0.33	0.12	0.20*	0.24*
	(0.08)	(0.15)	(0.33)	(0.12)	(0.16)	(0.22)	(0.10)	(0.09)	(0.10)
R^2	0.05	0.06	0.21	0.02	0.01	0.06	0.02	0.06	0.14
	Panel E: $\tilde{\phi}^{\mathrm{U}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\star}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{U}}_{t+h\|t}$								
α	0.62*	0.65*	0.65*	0.64*	0.73*	0.74*	0.58*	0.66*	0.71*
	(0.05)	(0.06)	(0.08)	(0.06)	(0.06)	(0.09)	(0.08)	(0.07)	(0.11)
β	0.16	0.20	0.53	0.07	0.12	0.46	0.13	0.11	0.27*
	(0.08)	(0.14)	(0.29)	(0.11)	(0.14)	(0.24)	(0.09)	(0.07)	(0.12)
R^2	0.04	0.04	0.13	0.01	0.01	0.08	0.04	0.03	0.14
	Panel F: $\tilde{\phi}^{\mathrm{B}}_{t+h\|t} = \alpha + \beta \times \tilde{f}^{\mathrm{B}}_{t+h\|t} + \tilde{\varepsilon}^{\mathrm{B}}_{t+h\|t}$								
α	0.58*	0.60*	0.61*	0.59*	0.72*	0.73*	0.52*	0.61*	0.71*
	(0.07)	(0.07)	(0.09)	(0.08)	(0.08)	(0.10)	(0.10)	(0.09)	(0.13)
β	0.17	0.27	0.67	0.10	0.11	0.57	0.15	0.14	0.29*
	(0.10)	(0.18)	(0.35)	(0.13)	(0.17)	(0.29)	(0.11)	(0.09)	(0.14)
R^2	0.03	0.04	0.14	0.02	0.01	0.09	0.03	0.03	0.13

Notes: This table displays the estimates of Eqns. (4.13) and (4.14) using only regular SPF participants and disagreement statistics based on the histogram means for each outcome variable and forecast horizons $h \in \{4, 8, 20\}$. The sample period $t = 1, \ldots, 78$ represents time instances between 1999Q1 and 2018Q2. Coefficients are estimated with OLS. Newey-West (1987) standard errors accounting for arbitrary levels of heteroskedasticity and autocorrelation are reported in parentheses. An asterisk '*' indicates significance at the 5% critical level against a two-sided alternative.

Table 4.12: Correlations of survey-based and other uncertainty proxies (robustness)

		Inflation rate			GDP growth rate			Unemployment rate		
Horizon h		4	8	20	4	8	20	4	8	20
EPU	$\overline{\sigma}^{\mathrm{M}}_{t+h\|t}$	0.74	0.75	0.67	0.69	0.72	0.68	0.72	0.72	0.65
	$\overline{\sigma}^{\mathrm{U}}_{t+h\|t}$	0.74	0.75	0.67	0.69	0.72	0.68	0.72	0.72	0.65
	$\overline{\sigma}^{\mathrm{B}}_{t+h\|t}$	0.74	0.75	0.67	0.70	0.71	0.68	0.71	0.71	0.65
	$\tilde{\phi}^{\mathrm{M}}_{t+h\|t}$	0.68	0.69	0.62	0.68	0.63	0.65	0.69	0.65	0.59
	$\tilde{\phi}^{\mathrm{U}}_{t+h\|t}$	0.72	0.72	0.62	0.59	0.62	0.59	0.68	0.66	0.62
	$\tilde{\phi}^{\mathrm{B}}_{t+h\|t}$	0.70	0.71	0.60	0.57	0.62	0.59	0.69	0.67	0.63
	$s_{t+h\|t}$	0.25	0.16	0.03	0.00	−0.09	0.19	0.17	0.17	0.22
	$s^{\star}_{t+h\|t}$	0.27	0.15	0.33	0.04	0.01	0.26	0.17	0.25	0.27
	$s^{\mathrm{B}}_{t+h\|t}$	0.27	0.14	0.31	0.04	0.00	0.27	0.17	0.25	0.27
	$\tilde{f}_{t+h\|t}$	0.08	0.09	−0.12	0.04	0.00	0.16	0.04	0.07	0.26
	$\tilde{f}^{\star}_{t+h\|t}$	0.12	0.12	0.46	0.08	0.06	−0.02	0.12	0.18	0.29
	$\tilde{f}^{\mathrm{B}}_{t+h\|t}$	0.14	0.09	0.47	0.08	0.04	−0.02	0.12	0.18	0.30
VSTOXX	$\overline{\sigma}^{\mathrm{M}}_{t+h\|t}$	0.02	0.03	−0.11	0.01	0.14	−0.01	−0.15	−0.11	−0.23
	$\overline{\sigma}^{\mathrm{U}}_{t+h\|t}$	0.01	0.03	−0.11	0.01	0.14	−0.01	−0.15	−0.11	−0.23
	$\overline{\sigma}^{\mathrm{B}}_{t+h\|t}$	0.01	0.04	−0.10	0.03	0.17	0.02	−0.14	−0.09	−0.22
	$\tilde{\phi}^{\mathrm{M}}_{t+h\|t}$	0.05	0.11	−0.10	0.04	0.18	0.09	−0.06	−0.05	−0.17
	$\tilde{\phi}^{\mathrm{U}}_{t+h\|t}$	0.02	0.04	−0.04	0.10	0.16	−0.03	−0.08	−0.07	−0.08
	$\tilde{\phi}^{\mathrm{B}}_{t+h\|t}$	0.04	0.07	−0.05	0.08	0.16	−0.01	−0.11	−0.09	−0.10
	$s_{t+h\|t}$	0.62	0.48	0.17	0.62	0.45	−0.27	0.61	0.54	0.29
	$s^{\star}_{t+h\|t}$	0.63	0.46	0.20	0.65	0.47	0.00	0.55	0.48	0.33
	$s^{\mathrm{B}}_{t+h\|t}$	0.63	0.47	0.20	0.65	0.47	−0.02	0.55	0.48	0.33
	$\tilde{f}_{t+h\|t}$	0.56	0.32	−0.01	0.58	0.37	−0.43	0.57	0.48	0.01
	$\tilde{f}^{\star}_{t+h\|t}$	0.52	0.35	−0.12	0.64	0.53	−0.14	0.47	0.40	0.09
	$\tilde{f}^{\mathrm{B}}_{t+h\|t}$	0.51	0.37	−0.12	0.64	0.53	−0.17	0.48	0.41	0.07

Notes: For each outcome variable and forecast horizon, this table displays the bivariate correlation between one particular measure of forecaster disagreement or survey-based uncertainty and either the Economic Policy Uncertainty (top panel) or the VSTOXX volatility indices (bottom panel). We calculate EPU and VSTOXX in quarter t as the average across all observations between the deadline for the previous survey round from quarter $t-1$ and the deadline for the current round (excluding the former but including the latter). The sample period is 1999Q1–2018Q2.

Chapter 5

Overconfidence Versus Rounding in Survey-Based Density Forecasts

5.1 Introduction

Forecasts that dispense with uncertainty bands are increasingly regarded as incomplete. It has been argued that to express how strongly a point prediction is expected to deviate from the ex-post observed outcome, point forecasts should be complemented by a quantification of ex-ante uncertainty (Dawid, 1984; Bruine de Bruin et al., 2010). While it has been documented that survey forecasts for inflation, GDP growth or unemployment outperform model-based forecasts (cf. Ang et al., 2007; Faust and Wright, 2009), the informative content of survey predictions for the conditional variance has been recently contested, e.g., by Clements (2016). In the case of the Survey of Professional Forecasters (SPF) that is conducted by the Federal Reserve Bank of Philadelphia (FED) and the European Central Bank (ECB), point forecasts are elicited along with probabilistic forecasts in the form of histograms. This allows to derive a measure of ex-ante uncertainty by computing the variance of the reported histograms. Several desirable properties of this index have been documented. For example, Lahiri and Sheng (2010) document that the cross-sectional average variance increases with the forecast horizon. However, it has been found that the ex-ante variance (in our terms, 'uncertainty') deviates considerably from the average squared ex-post forecast error. This is sometimes referred to as 'over- or underconfidence' (Kenny et al., 2014, 2015; Clements, 2014). The term 'overconfidence' might in this context either be understood to reflect the inherent characteristics of forecasters or rather as a mere description of an ex-ante variance that is small compared

This chapter is based on a working paper of the same name that I wrote jointly with Matthias Hartmann (cf. Glas and Hartmann, 2018).

to a predefined benchmark such as the ex-post variance. However, this finding suggests that the average variance of the SPF histograms as proposed by Zarnowitz and Lambros (1987) has to be interpreted cautiously.

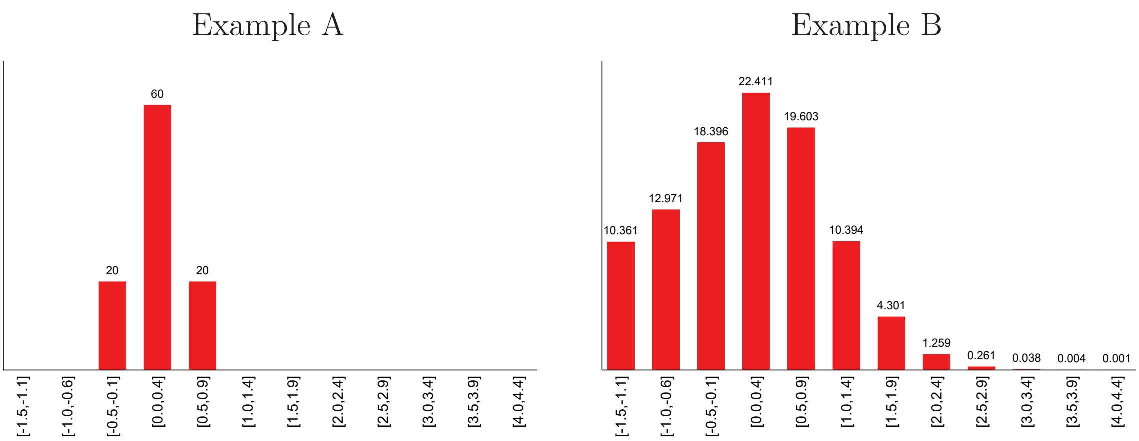
In this chapter, we ask under which conditions the second moments from the SPF data are relatively well aligned with the variability of the prediction errors. The derivation of an ex-ante measure of forecast uncertainty that takes potential distortions into account is difficult since the survey data does not contain any covariates that might help to understand forecasters' behavior.[1] Thus, hypotheses about the dependence of individuals' reported ex-ante uncertainty on misperceptions of their own capability to forecast cannot be easily examined empirically. Given these difficulties, we propose to relate the ex-ante variance of forecasters to the properties of the predictions itself which are observed prior to the outcome. These characteristics are the distinct rounding patterns in individual histogram forecasts and the differences in the survey design that are laid down in the regulations of the U.S. FED on the one hand and the ECB on the other hand. It has been documented in several empirical and experimental studies that (i) the individual responses in many surveys, including the SPF, are subject to considerable rounding (Manski and Molinari, 2010) and that (ii) individual uncertainty and forecasters' tendency to round that is observed in survey responses are related (Boero et al., 2011; Ruud et al., 2014; Binder, 2017). Moreover, according to Manski and Molinari (2010), the extent of rounding varies across individual survey participants. The same cross-sectional variation in the SPF-based uncertainty has been frequently documented since Zarnowitz and Lambros (1987).

Empirically, a misalignment of ex-ante and ex-post forecast variances has been described by Giordani and Söderlind (2003, 2006), Kenny et al. (2014, 2015) and Clements (2014). Our main finding is that the deviations of survey participants' forecast variances prior to and after the outcome can partially be ascribed to the response pattern of a large group of forecasters that provide their histogram predictions in a particular form. A striking feature of this group is that their forecasts are conveyed in a rather coarse form, with apparently strongly rounded numbers and a relatively low number of probability categories that are assigned nonzero numbers. An example of this is depicted in Figure 5.1. The subfigures show histogram forecasts for the annual inflation rate in 2016 reported by two participants in the 2016Q4 survey wave of the ECB-SPF. Two differences are apparent. First, the forecasted probabilities in Example A are multiples of 10%, whereas those in Example B do not seem to have a common divisor. Second, the number

[1]One notable exception is an indicator variable in the FED-SPF data that reports whether forecasters are employed in the financial services industry, a research institute or any other employer.

of outcome intervals that contain nonzero probability numbers is considerably smaller in the left graph. In other words, the right histogram exhibits larger variance.

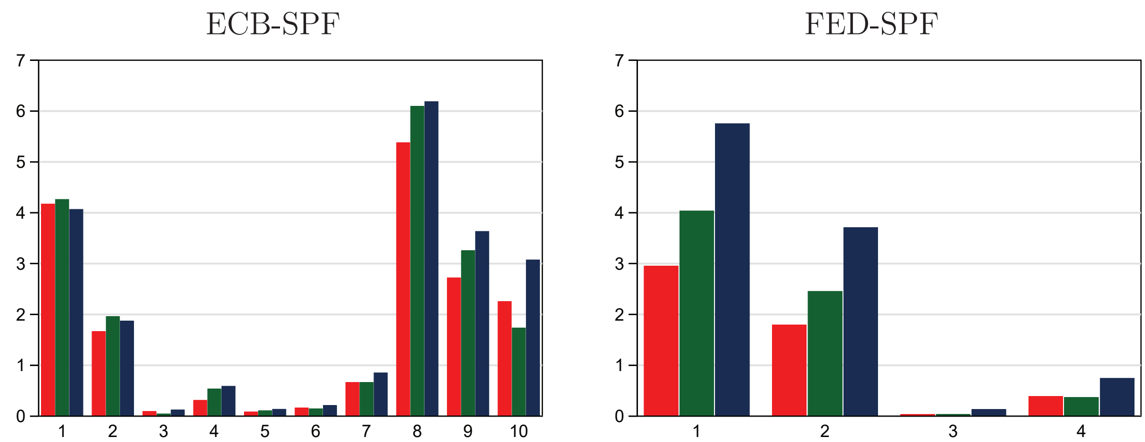
Figure 5.1: Two examples of histogram forecasts from the ECB-SPF



Notes: The graphs depict two examples of one-quarter-ahead histogram forecasts for the inflation rate from the ECB-SPF. Both predictions are taken from the 2016Q4 survey. The left plot depicts the histogram of forecaster 102. The right subfigure displays the probabilities reported by forecaster 95. In the latter case, the decimal numbers attached to the probabilities are cut off at the third decimal, i.e., the original histogram in the SPF data contains additional decimal numbers.

Moreover, Figure 5.2 summarizes the share of probabilities that contain between one and ten decimal numbers out of all reported probability numbers in the SPF data, pooled across forecasters, time instances and prediction horizons. The '0'-category is omitted to improve the readability. The left panel of Figure 5.2 shows that the ECB-SPF contains two clearly separated groups of forecasts that are distinguished in terms of the number of digits in their numerical values. The right part of the figure shows the counterpart for the case of the FED-SPF. As it is suggested in Figure 5.1, separating the two groups, we find that the ex-ante variances of those forecasters who report more strongly rounded numbers are substantially smaller than those of the survey participants who appear to round less or not at all. Moreover, the ex-ante and ex-post variances of the non-rounding group of forecasters are clearly more in line with each other than in the case of the group which reports strongly rounded histogram probabilities. This holds for both the ECB- and the FED-SPF. However, the number of responses that entail a large number of digits is substantially larger in the former than in the latter. This observation provides an important hint regarding a potential explanation for the rounding choice of survey participants. The FED-SPF is elicited in a rather traditional way by asking respondents to fill in the questionnaire by means of paper-and-pencil. In contrast, the ECB-SPF questions can be answered on the computer via an Excel spreadsheet. Hence, it is probably the case that respondents find it substantially easier to report numbers with many digits in the case of the ECB-SPF.

Figure 5.2: Relative frequencies of the number of decimals reported in the SPF



Notes: The graphs depict the share of probabilities that contain $d \in \{1, \ldots, 10\}$ decimal numbers out of all reported probabilities in the SPF data for inflation, output growth and unemployment. The '0'-category is omitted to improve the readability. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Naturally, other explanations are likely to play a role besides this hypothesis. In Section 5.6, we describe empirical evidence from two special surveys conducted by the ECB-SPF which indicate that two groups of survey participants can be separated based on the degree to which they rely either on formal models or judgment when producing their forecasts. Interestingly, the size of these groups roughly corresponds to the one of the rounders and non-rounders that we find based on distinct classification schemes. However, an analysis of a direct connection between rounding and such (non-)modeling choices at the individual forecaster level is infeasible since the results of the special question are reported without identification numbers. This means that while it provides an intuitive explanation, the relation between modeling choices and rounding is essentially untestable based on the available data.

Our findings have three practical implications. First, the distortion of an index of overall uncertainty that is computed as the average across the individual variances (Lahiri and Sheng, 2010; Lahiri et al., 2015) can be reduced ex-ante by discarding those forecasts that are strongly rounded. Second, a more general improvement in the variance forecasts of the FED-SPF may be achieved if survey participants were given the opportunity to fill in the questionnaire online. Third, the trajectory of average uncertainty during recent years is at least partly affected by the overall increase in the share of forecasters that do not report strongly rounded numbers. Thus, we conclude that while uncertainty is likely higher than what is reflected by the average forecast variance across all forecasters due to the presence of considerable rounding, the increase in uncertainty during the years after the crisis is likely overstated due to changes in the composition of the SPF panel.

The remainder of this chapter is structured as follows. After briefly reviewing the related literature in Section 5.2, the data are introduced in Section 5.3. Next, we discuss the categorizations that are used to classify the survey participants as rounders or non-rounders in Section 5.4. The findings regarding the forecast performance of the density predictions are presented in Section 5.5. A detailed discussion of the findings follows in Section 5.6. Section 5.7 summarizes and concludes.

5.2 Rounding and the Information Content of Histogram Forecasts

Among many ways to quantify the uncertainty of forecasts, surveys like the SPF have become one particularly popular means during recent times. However, it is still not well understood to what extent numerical inaccuracies such as rounded numbers may distort the variance of histogram forecasts. In a fairly general framework, Heitjan and Rubin (1991) discuss the implications of rounding and similar forms of incomplete survey responses on the likelihood of parameter estimates that are based on survey data. Similarly, Tay and Wallis (2002) note that the communication of uncertainty by means of survey-based density forecasts faces several distinct problems. Some of the crucial steps, like the design of the survey questionnaire, the timing of the elicitation process, the production and reporting of forecasts by survey participants as well as the interpretation and evaluation by users of the survey may introduce distortions in the conveyed information. Perhaps owing to such difficulties, various objections towards the use of survey data in economics in general have been formulated, some of which are summarized by Manski and Molinari (2010).

The central question of this chapter is how rounding may affect ex-ante and ex-post measures of forecast variance. We are particularly interested in the implications of the observation that forecasters who provide strongly rounded responses also show a striking tendency to provide narrow histograms with only a small number of outcomes to which they attach nonzero probabilities. It has been previously noted that such response behavior may affect conditional second moment statistics from survey data. For example, Boero et al. (2015) interpret the decision of forecasters to round the probabilities of histograms in surveys as an expression of what they call “uncertain uncertainty”. Other studies such as Manski and Molinari (2010) also highlight the importance of rounding choices on the outcomes of histogram forecasts as they are provided by the SPF.

A distinct approach is taken by Binder (2017), who derives an index of uncertainty based on rounding outcomes in a survey of inflation expectations. The construction of the index in Binder (2017) is based on the assumption that rounding can be seen as an expression of uncertainty. This is also reflected in Bruine de Bruin and Carman (2012) or Ruud et al. (2014). These hypotheses regarding the link between rounding and uncertainty connect to the more general literature which discusses rounding and other forms of *data coarsening* (Heitjan and Rubin, 1991; Ruud et al., 2014). In contrast to the approach of Binder (2017), we do not employ rounding as the single source of information regarding uncertainty, but derive a direct measure of uncertainty based on the SPF histograms. This enables us to discuss potential distortions from rounding in the computation of the resulting uncertainty index.

In a recent paper, Clements (2016) examines the informative content of density forecasts in terms of their capability to deliver variance forecasts and concludes that the SPF data provided by the ECB contains little reliable information beyond the forecast for the conditional mean. In the current study, we draw upon such findings and examine to what extent they can be understood as the result of data coarsening in the form of rounding and the tendency to concentrate the entire probability mass in a small share of the outcome intervals from the survey questionnaire. In a related study, Clements (2011) documents that the mismatch between the reported probabilities of a decline in output growth and corresponding probabilities derived from the histogram forecasts can be partially explained by the rounding choices of the forecasters in the FED-SPF. Since more than 75% of the SPF participants' responses appear to be rounded to some extent, it seems recommendable to investigate the implications of this particular data feature for the assessment of macroeconomic uncertainty.

5.3 Data

In this section, the data used to quantify ex-ante uncertainty in both the Euro area and the U.S. are described.

The survey data are provided by the SPF of the ECB and the U.S.-FED. Both surveys elicit point and density forecasts of future inflation, real GDP growth and unemployment rates in the Eurozone and the U.S. at the quarterly frequency.[2] For inflation and output growth, the outcome variable x_t refers to year-on-year growth rates, i.e.,

$$x_t = 100 \times \left(\frac{X_t}{X_{t-1}} - 1 \right), \tag{5.1}$$

[2]We use the terms 'density forecast' and 'histogram forecast' synonymously throughout this chapter.

where X_t denotes the annual average of either the respective price index or real GDP in year $t = 1, \ldots, T$.[3] In the case of the unemployment rate, x_t is calculated as the annual average over the civilian unemployment rates that are observed at the monthly frequency (i.e., $x_t = X_t$). Data on the realizations for the Euro area and the U.S. are drawn from the Statistical Data Warehouse of the ECB and the Real-Time Data Set for Macroeconomists of the Federal Reserve Bank of Philadelphia, respectively. Both databases provide data vintages for all outcome variables. For each vintage, we calculate X_t in all cases where consecutive observations for each month (Harmonized Index of Consumer Prices, unemployment rate) or quarter (GDP price index, real GDP) of year t are provided and compute x_t. In the empirical analysis, we employ the first-releases of x_t, which are most closely related to the information available to forecasters when they produce their predictions. Moreover, Jo and Sekkel (2018) show that ex-post forecast variances based on the most recent data vintage tend to be underestimated.

The survey data from the SPF consist of so-called 'fixed-event' density forecasts, which are characterized by a fixed target year t and a rolling quarterly forecast horizon h. The nature of these forecasts implies that h diminishes in each consecutive quarter in which the survey is conducted until the arrival of the realization in t. We consider both the predictions for the current and the next year. This obtains a sequence of individual h-step-ahead density forecasts with forecast horizons $h \in \{8, 7, \ldots, 1\}$. In the case of the inflation rate and output growth, forecasters in our sample target the years 2000 to 2016. This means that the time period when forecasts are made and collected ranges from 1999Q1 to 2016Q4.[4] Density forecasts for the unemployment rate in the FED-SPF are available only since 2009Q2, whereas the responses in the ECB-SPF are available for the entire sample period. For the U.S., we thus focus on the unemployment rates in the years 2011 to 2016, for which h-step-ahead predictions are available for each forecast horizon.

In the questionnaire, survey participants $i = 1, \ldots, N$ are requested to assign probabilities to a prespecified number of outcome intervals (so-called 'bins'). Let $p_{i,k,t,h} \in [0, 100]$ for $k = 1, \ldots, K$ denote the probability number assigned to the k-th bin. The bins have

[3]The ECB-SPF inflation forecasts refer to the monthly Harmonized Index of Consumer prices. For the FED-SPF we use the quarterly chain-weighted GDP price index. We prefer GDP inflation over CPI inflation because density forecasts for the latter are only available since 2007 in the FED-SPF, whereas predictions for the former are available for the entire sample period. For the computation of output growth, we use quarterly real GDP.

[4]Forecasts for inflation and output growth in the U.S. are available since 1968Q4. However, we prefer to focus on a common sample period and exclude these earlier predictions. This also helps to avoid various methodological changes in the FED-SPF such as the switch from gross national product to gross domestic product. Since no five- to eight-step-ahead forecasts for the year 1999 are available in the ECB-SPF data, we exclude the current year predictions from the surveys conducted between 1999Q1 to 1999Q4. Similarly, no one- to four-step-ahead forecasts for the year 2017 are available in both surveys.

a width of 0.4 percentage points in case of the ECB-SPF as can be seen in Figure 5.1. In the case of the FED-SPF, the bin width is 0.9 percentage points except in a few cases.[5] As in Abel et al. (2016), the gaps between the interior bins are closed by extending the lower and upper bound of each bin by 0.05 percentage points. This seems to be in line with how most of the survey participants interpret their reporting task, as it is documented in a special survey conducted by the ECB in 2008, where 76% of the SPF participants stated that they interpret an interval like $[1.5, 1.9]$ to actually indicate a range as given by $[1.45, 1.95]$ (cf. ECB, 2009). The bins at the lower and upper end of the support are assumed to have twice the width of the interior intervals, i.e., one or two percentage points depending on the survey and variable. The bounds of the individual histograms are fixed at the leftmost and rightmost bin with nonzero probability mass. Moreover, the maximum range covered by the bins differs across surveys, outcome variables and time instances.

We exclude observations from the sample whenever the sum over the reported probabilities deviates by at least 0.9 percentage points from the required 100% overall probability in absolute terms.[6] Moreover, we delete all histograms that contain 100% probability in a single bin. Based on the approach used to compute moments of the histogram forecasts that is introduced below, such cases would result in a variance of zero. This is likely to affect our conclusions disproportionately. However, such responses are relatively scarce in both surveys.[7]

The participants in both surveys include employees of research institutes and the financial services industry. The occupation of the anonymous survey participants is provided in the case of the FED-SPF. Depending on the survey period under consideration, 22–50% of the participants of the FED-SPF are classified as 'financial service providers' and 39–70% as 'non-financial service providers'. A third category of unclassified 'others' is also included, which amounts to 0–15% of the cross-section. In the case of the ECB-SPF, this information is not provided. An identification number allows to track the anonymous individual forecasters. We observe a relatively large number of entries and exits of SPF participants in each survey round. In order to analyze whether participation varies systematically across the different forecast horizons, we define the participation indicator variable $D^{\mathcal{P}}_{i,t,h}$, which is equal to unity if forecaster i issues an h-step-ahead density forecast for x_t, and zero else. For each forecast horizon $h \in \{8, 7, \ldots, 1\}$, Table 5.1 displays the

[5] Since 2014Q1, the bin width for inflation is 0.4 percentage points. Similarly, the interior bins for the unemployment rate have a width of 0.4 percentage points throughout the sample period.

[6] We permit small deviations in order to keep the non-rounded histograms in the sample. In such cases the probabilities may not add up to exactly 100%.

[7] Less than 1% of the histogram forecasts reported in the ECB-SPF and around 2% in the FED-SPF.

number of density forecasts reported in both versions of the SPF, i.e., $\sum_{t=1}^{T}\sum_{i=1}^{N} D_{i,t,h}^{\mathcal{P}}$.

Table 5.1: Number of density forecasts provided by SPF participants

		Forecast horizon h								
SPF	Variable	8	7	6	5	4	3	2	1	$\sum_h$
ECB	Inflation	894	905	814	919	910	900	803	814	6959
	GDP growth	902	909	818	928	919	909	812	885	7082
	Unemployment	864	869	779	869	869	861	760	798	6669
FED	Inflation	590	612	594	621	614	613	572	525	4741
	GDP growth	615	638	615	646	638	645	604	577	4978
	Unemployment	228	219	222	229	221	217	209	187	1732

Notes: For each outcome variable, this table displays the number of reported histograms per forecast horizon, i.e., $\sum_i \sum_t D_{i,t,h}^{\mathcal{P}}$, as well as the total number of observations across all horizons. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

The sample size is roughly constant across variables and forecast horizons in both surveys with the obvious exception of the unemployment rate in the FED-SPF. This suggests that the cross-section of forecasters is relatively similar. Although the total number of participants is larger in the FED- than in the ECB-SPF (115 versus 102), the sample size for both inflation and real GDP growth is considerably larger in the latter case. In other words, average participation is lower in the FED-SPF. Between 1999Q1 and 2016Q4, the number of forecasters who contributed to the ECB-SPF has declined from 63 to 39. Over the same time period, the number of participants in the FED-SPF is relatively constant. In particular, 28 forecasters have submitted predictions in 1999Q1 compared to 32 in 2016Q4.

In order to compute first and second moments of the histograms, it is common to assume that the entire probability mass within each bin is located at the midpoint. Alternatively, one may compute the moments of a smoothed density function as it is done in Engelberg et al. (2009) or Glas and Hartmann (2016). However, the imputation of a continuous density precludes the analysis of rounding choices and is, thus, not suitable for our purposes. Based on the 'mass-at-midpoint' approach, the mean of forecaster i's histogram is given by

$$\mu_{i,t,h} = \frac{1}{100}\sum_{k=1}^{K} p_{i,k,t,h} \times m_k, \tag{5.2}$$

with m_k denoting the midpoint of the k-th bin. The h-step-ahead 'consensus' forecast is calculated as the equally-weighted average over the individual histogram means, that is,

$$\bar{\mu}_{t,h} = \frac{1}{N} \sum_{i=1}^{N} \mu_{i,t,h}. \tag{5.3}$$

In order to analyze which data release is predicted by the SPF participants, Figure 5.3 depicts the realizations of each outcome variable in the Euro area and the U.S. using observations from both the first release (solid line) and the most recent data vintage (dashed line). Moreover, each plot includes the consensus forecasts, i.e., $\bar{\mu}_{t,h}$ from Eqn. (5.3), for horizons $h \in \{8, 7, \ldots, 1\}$. The one- and eight-step-ahead predictions are highlighted distinctly from the other forecast horizons.

The evidence from Figure 5.3 shows that the accuracy of the average forecast improves as the target period approaches. In other words, forecast errors decline with h. In particular, the deviation between x_t and $\bar{\mu}_{t,1}$ is smaller than the difference between x_t and $\bar{\mu}_{t,8}$ in almost all cases. Moreover, in cases where the first and last releases of the data deviate substantially, $\bar{\mu}_{t,1}$ is more closely associated with the former. This finding suggests that participants of the SPF predict the first release of the respective outcome variable. This supports our choice of focusing on this particular data release in the empirical analysis. However, we have also used last-release data in a robustness check, which has little impact on the empirical findings.

To compare the mismatch between ex-ante and ex-post uncertainty, we need a quantification of the variances of the reported histograms that enables us to retain the information regarding the rounding choices of forecasters. Based on the means from Eqn. (5.2), we calculate the individual variance as

$$\sigma^2_{i,t,h} = \frac{1}{100} \sum_{k=1}^{K} p_{i,k,t,h} \times (m_k - \mu_{i,t,h})^2. \tag{5.4}$$

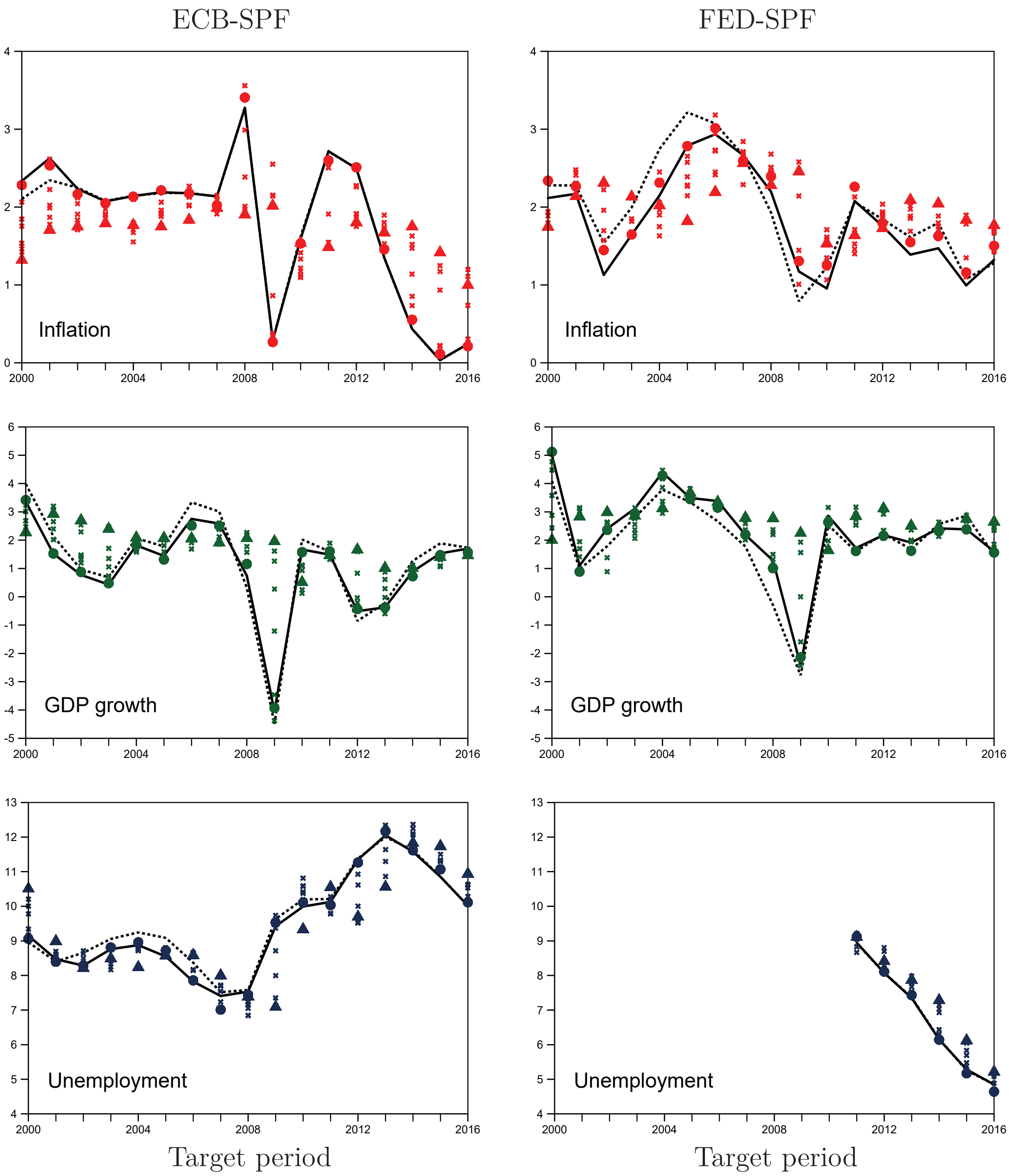
This variable serves as a measure of forecaster i's ex-ante uncertainty. To obtain an indicator of aggregate uncertainty, we follow Lahiri and Sheng (2010) and compute the cross-sectional average of the h-step-ahead variances from Eqn. (5.4),

$$\overline{\sigma^2_{t,h}} = \frac{1}{N} \sum_{i=1}^{N} \sigma^2_{i,t,h}. \tag{5.5}$$

Analogously to Figure 5.3, Figure 5.4 depicts the time series of the h-step-ahead average forecast variances.

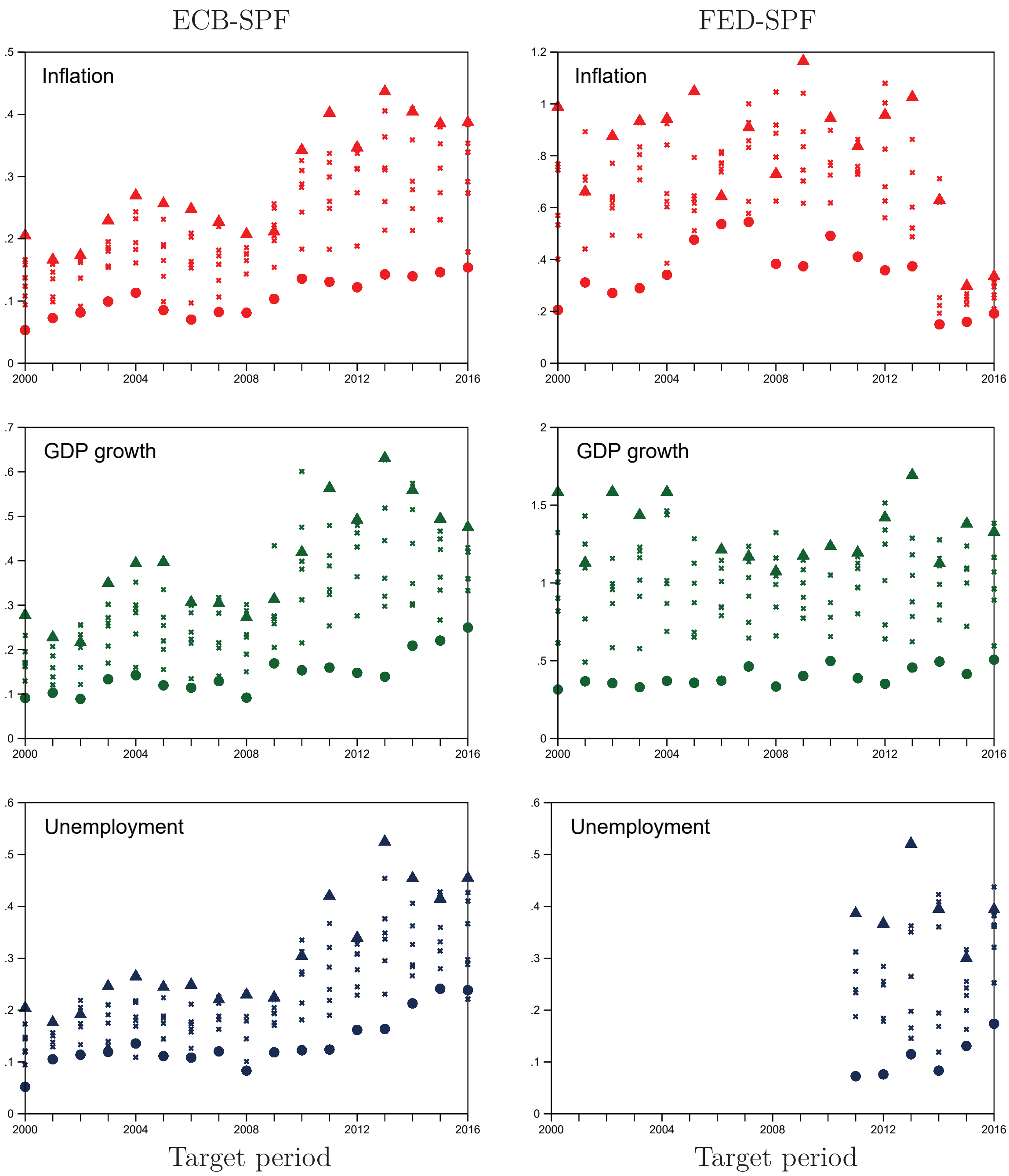
Average ex-ante uncertainty declines with the forecast horizon, i.e., the average forecaster becomes increasingly more confident as the target period approaches and more

Figure 5.3: Realizations and consensus forecasts from the SPF



Notes: The graphs depict the time series of the annual realizations for inflation (first row), output growth (second row) and unemployment (third row) in the Eurozone and the U.S. based on first-release (solid black lines) and last-release (dashed black lines) data vintages. In addition, each plot displays the cross-sectional average across the means of the individual h-step-ahead histogram forecasts, i.e., $\bar{\mu}_{t,h}$ from Eqn. (5.3). Triangles '△' and bullets '•' indicate the eight- and one-step-ahead consensus forecasts, respectively. Crosses '×' indicate the predictions for the intermediate forecast horizons. The horizontal axis depicts the target year. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.4: Average ex-ante uncertainty



Notes: The graphs depict the time series of the cross-sectional average across the h-step-ahead variances from the individual histograms for inflation (first row), output growth (second row) and unemployment (third row) in the Eurozone and the U.S., i.e., $\overline{\sigma^2_{t,h}}$ from Eqn. (5.5). Triangles '△' and bullets '•' indicate the eight- and one-step-ahead average variances, respectively. Crosses '×' indicate the average variances for the intermediate forecast horizons. The horizontal axis depicts the target year. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

information about the realization is available. Moreover, an increase in average uncertainty is visible in most cases after the outbreak of the financial crisis in 2008. Owing to an adjustment of the bin definitions in 2014Q1, a break in the time series of ex-ante uncertainty is visible for the predictions of the inflation rate in case of the FED-SPF. However, there is almost no effect on our results if the data for the years 2014-2017 is discarded. Thus, we decided to keep these observations in the sample.

So far, we have described the characteristics of the entire cross-section of SPF participants in both the U.S. and the Euro area. However, it may be that panelists differ systematically with respect to the coarseness of their predictions. In the next step, we aim to isolate two distinct groups of forecasters based on the way that individual survey participants decide to round (or not to round) the reported probability numbers.

5.4 Rounding Schemes

In this section, we discuss alternative classification schemes that serve as a means to distinguish non-rounders from rounders based on their reporting behavior.

Though rounding is one of the most striking characteristics of the histogram forecasts in the SPF, an unambiguous classification into rounders and non-rounders is not possible. Since the coarseness of the responses appears to vary across individual forecasters, we propose several distinct categorization schemes in order to assess the robustness of our findings. Due to the anonymous nature of participation in the SPF, reputational concerns should not play an important role in the decision whether or not to round a prediction. In most empirical research on rounding of survey-based forecasts, the participants are classified as rounders based on whether the point forecast is a multiple of an integer number (e.g., Binder, 2017). In contrast, we analyze the histograms reported in the SPF. Thus, the employed rounding schemes are based on multiple reported numbers for each individual, instead of just a single one. Moreover, we consider two distinct types of categorizations that differ in terms of what constitutes a rounded probability.

5.4.1 Decimal-Based Categorization

The first type of categorization is based on the number of decimals of each probability number, $p_{i,k,t,h}$, which is denoted as $d_{i,k,t,h}$. For notational convenience, we suppress all subscripts except for i and k in the following subsections.[8] Let $K_i \in \{2, \ldots, K\}$ denote the number of bins to which forecaster i assigns nonzero probability, i.e., cases where $p_{i,k} > 0$.

[8]This does not mean that we assume that variation across time or forecast horizons plays no role. We analyze this in Section 5.5.

Similarly, $K_i^\star \in \{0, \ldots, K_i\}$ indicates the number of bins with nonzero probability that contain decimals numbers, i.e., cases where both $p_{i,k} > 0$ and $d_{i,k} > 0$. The share of probabilities in forecaster i's histogram that contain nonzero decimals numbers is thus given by

$$\rho_i = \frac{K_i^\star}{K_i}. \tag{5.6}$$

Based on ρ_i, we define distinct classification schemes that are introduced here in terms of how strictly we delineate the definition of a non-rounder. That is, each of the rules that are successively introduced below is less likely to classify a forecaster as a non-rounder than the previous one. The first approach is to treat a forecaster as a non-rounder if *any* of the individually reported probability numbers are stated by means of using decimals, that is,

$$D_i^{\text{any}} = \begin{cases} 1 & \text{if } \rho_i > 0 \text{ and} \\ 0 & \text{else.} \end{cases} \tag{5.7}$$

It is likely that this indicator will classify some forecasters as non-rounders even though the majority of reported numbers entail a rather strong degree of rounding. Consider an example where five bins are available, i.e., $K = 5$, and a survey participant reports probabilities $(p_{i,1}, \ldots, p_{i,5})' = (0.5\%, 30\%, 39\%, 30\%, 0.5\%)'$, such that $K_i = 5$ and $K_i^\star = 2$. Despite the fact that only the probabilities in the tails include decimals, such a forecaster is considered as a non-rounder based on D_i^{any} since $\rho_i = 0.4$. A more restrictive rule to single out non-rounders is obtained if a survey participant is regarded as a non-rounder if *most* of the probabilities are reported with nonzero decimal numbers, i.e.,

$$D_i^{\text{most}} = \begin{cases} 1 & \text{if } \rho_i > 0.5 \text{ and} \\ 0 & \text{else.} \end{cases} \tag{5.8}$$

This approach categorizes forecasters as non-rounders if more than 50% of the probabilities reported in a given histogram contain decimal numbers. Note that if K_i is even and half of the probabilities contain decimals while the other half do not, i.e., if $K_i^\star = K_i/2$, the scheme in Eqn. (5.8) classifies a survey participant as a rounder. Based on this categorization, the forecaster from the example above is considered to be a rounder because only 40% of the probabilities contain decimal numbers. The most restrictive approach is to classify a forecaster as a non-rounder if *all* probabilities are stated by means of nonzero decimal numbers, that is,

$$D_i^{\text{all}} = \begin{cases} 1 & \text{if } \rho_i = 1 \text{ and} \\ 0 & \text{else.} \end{cases} \tag{5.9}$$

In this case, forecasters are only considered to be non-rounders if each probability number is stated with nonzero decimal numbers, i.e., cases where $K_i^\star = K_i$. Based on the scheme in Eqn. (5.9), the forecaster from the example above is considered as a rounder because three out of five probabilities do not contain decimal numbers.

To summarize, the categorizations described in Eqns. (5.7)-(5.9) classify survey participants as rounders if any, most, or all of the probabilities are stated with nonzero decimal numbers. It thus follows that $\sum_{i=1}^N D_i^{\text{any}} \geq \sum_{i=1}^N D_i^{\text{most}} \geq \sum_{i=1}^N D_i^{\text{all}}$.

5.4.2 Integer-Based Categorization

Binder (2017) classifies consumers as rounders based on whether their point forecast is a multiple of five. Similarly, Manski (2004) notes that probabilistic forecasts are frequently multiples of an integer number. For example, D'Amico and Orphanides (2006), Engelberg et al. (2009) or Clements (2011) observe that the probabilities reported in the FED-SPF tend to be multiples of five or ten. Boero et al. (2015) documents similar evidence for the predictions from the Survey of External Forecasters. A similar integer-based approach is considered here, which contrasts with the previous categorization that classifies survey participants based on whether the reported probabilities contain decimal numbers. In order to analyze whether the decimal- and integer-based approaches yield comparable results in isolating rounders and non-rounders, we analyze whether the probability in the k-th bin of forecaster i's histogram is a multiple of integer $\tau \in \mathbb{N}$ by defining

$$\tilde{D}_{i,k}^{\text{m}\tau} = \begin{cases} 1 & \text{if } \tau \cdot \left\lfloor \frac{p_{i,k}}{\tau} \right\rfloor = p_{i,k} \;\; \text{and} \\ 0 & \text{else,} \end{cases} \tag{5.10}$$

where $\lfloor p_{i,k}/\tau \rfloor$ is the integer part of $p_{i,k}/\tau$. Based on the bin-specific indicator variables from Eqn. (5.10), forecasters are classified as rounders according to the following rule:

$$\tilde{D}_i^{\text{m}\tau} = \begin{cases} 1 & \text{if } \text{mode}(\tilde{D}_{i,1}^{\text{m}\tau}, \ldots, \tilde{D}_{i,K}^{\text{m}\tau}) = 1 \;\; \text{and} \\ 0 & \text{else.} \end{cases} \tag{5.11}$$

Thus, a survey participant is treated as a rounder if the majority of the probabilities are multiples of τ. If the modal value in Eqn. (5.11) is not uniquely defined, we set $\tilde{D}_i^{\mathrm{m}\tau}$ to zero. Thus, if half of the probabilities are multiples of τ, but the other half are not, the corresponding forecaster is considered a non-rounder. Note that $\tilde{D}_i^{\mathrm{m}\tau}$ is used to isolate rounders, whereas the decimal-based categorizations in Eqns. (5.7)-(5.9) isolate non-rounders. In order to faciliate the comparison between both approaches, we use

$$D_i^{\mathrm{m}\tau} = 1 - \tilde{D}_i^{\mathrm{m}\tau} \tag{5.12}$$

in most cases instead of $\tilde{D}_i^{\mathrm{m}\tau}$. Thus, forecasters are considered to be non-rounders if most of the probabilities are *not* multiples of τ. In reference to the evidence documented in Boero et al. (2015) that many of the probabilities submitted to the SPF are multiples of five or ten, the forecaster considered in the example from the previous subsection is classified as a non-rounder based on both $D_i^{\mathrm{m}5}$ and $D_i^{\mathrm{m}10}$ since only two out of the five probabilities are multiples of either five or ten.

5.5 Empirical Analysis

In this section, we report descriptive statistics for the groups of rounders and non-rounders and examine if this distinction helps to understand the finding of a mismatch between the ex-ante and ex-post variances of individual forecasters that has been repeatedly documented in the empirical literature. In order to test if the variance misalignment and rounding choices are systematically related, inferential results regarding the differences in the histogram characteristics of rounders and non-rounders are reported.

5.5.1 Descriptive Analysis of Rounders and Non-Rounders

Based on the methodology discussed in Section 5.4, we analyze the size and characteristics of the groups of rounders and non-rounders in the SPF data. The results from the empirical analysis are robust to the choice of the considered categorization. For the sake of brevity, we choose to focus on one of the decimal-based categorizations and consider the integer-based approach for one particular value of τ in the following sections.

Figure 5.2 shows that relatively few participants in the FED-SPF state their probabilities in terms of decimal numbers. In contrast, the share of probabilities that contain decimal numbers is considerably larger in the ECB-SPF. Moreover, the participants in the FED-SPF use a relatively narrow range of at most four decimals, whereas the panelists in the ECB-SPF use up to ten. This may be due to systematic differences in either the cross-section or the structure of both surveys such as the differences in the bin width.

Based on the small number of probabilities with $d_{i,k,t,h} > 0$ in the case of the FED-SPF, we choose to focus on $D^{\text{any}}_{i,t,h}$ as the preferred decimal-based classification scheme. This is recommendable since the explanatory power of the distinction between rounders and non-rounders may be reduced due to the smaller number of forecasters that are classified as non-rounders based on $D^{\text{most}}_{i,t,h}$ and $D^{\text{all}}_{i,t,h}$.

The choice of τ for the integer-based categorization is guided by the evidence from Figure 5.5, which depicts the share of rounded histograms in the SPF data based on Eqn. (5.11) for a pooled sample of observations across all forecast horizons in the cases of the inflation (first row), output growth (second row) and unemployment rates (third row). This share is calculated as 100 times the number of rounded histograms that are classified by means of $\tilde{D}^{\text{m}\tau}_{i,t,h}$ for $\tau \in \{1, \ldots, 10\}$ divided by the total number of reported predictions, i.e.,

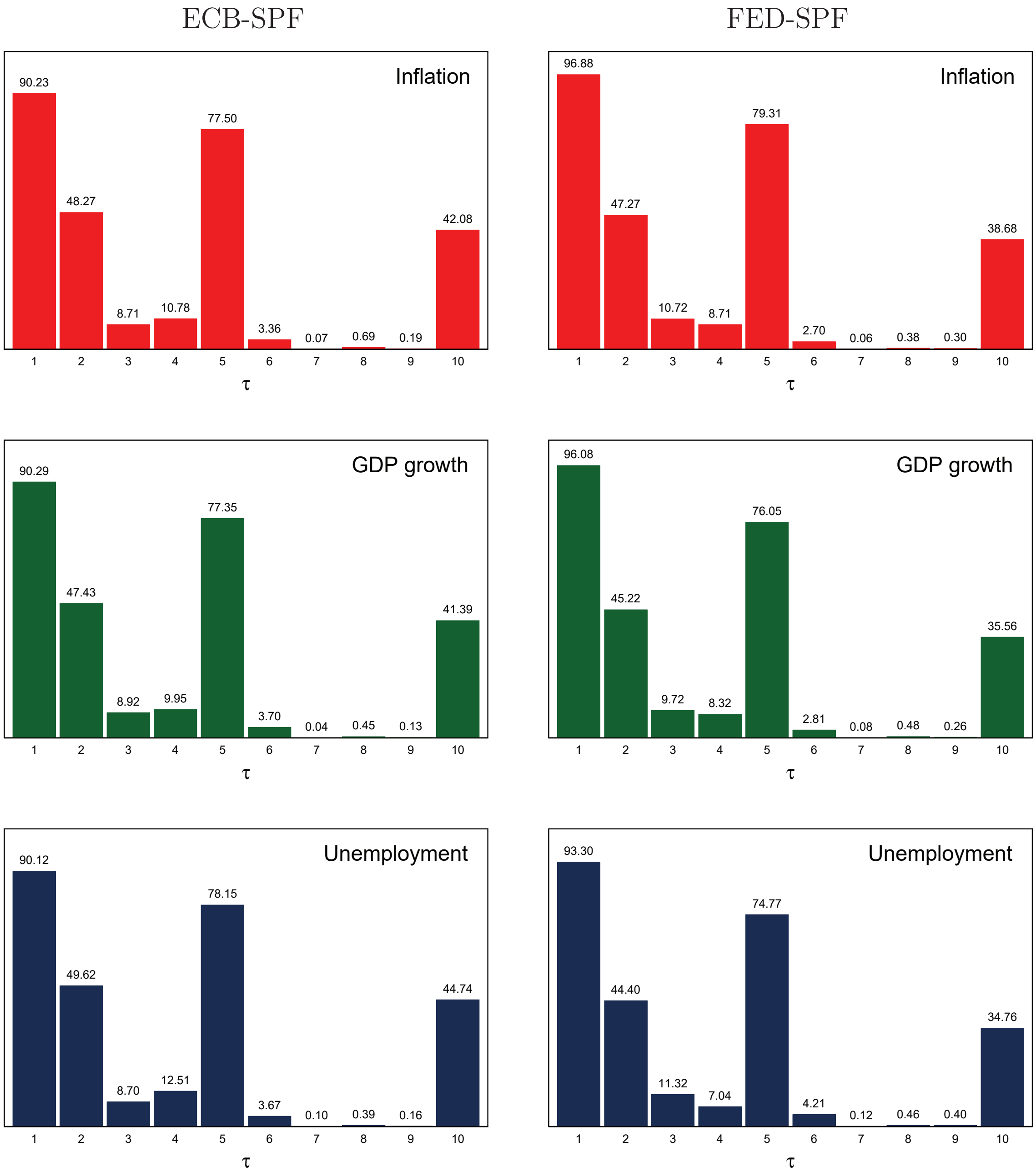
$$\tilde{S}^{\text{m}\tau} = 100 \times \frac{\sum_i \sum_t \sum_h \tilde{D}^{\text{m}\tau}_{i,t,h}}{\sum_i \sum_t \sum_h D^{\mathcal{P}}_{i,t,h}}. \tag{5.13}$$

The results are remarkably similar across all outcome variables and both versions of the SPF. The majority of survey participants are classified as rounders if we set τ to unity, i.e., many of the histograms are made up of probabilities that almost exclusively do not contain decimal numbers. This is not surprising given that Figure 5.2 shows that only a small fraction of the SPF participants reports probabilities with decimal numbers. There are two notable spikes in the cases where τ is set to either five or ten. This squares with the evidence documented in Engelberg et al. (2009) and Boero et al. (2015), who show that many of the probabilities reported in surveys of macroeconomic expectations are multiples of five or ten. In particular, 75–79% of all histograms in the SPF data consist of probabilities that are for the most part multiples of five. Similar numbers are reported in Clements (2011). Thus, we isolate non-rounders by setting τ to five and use $D^{\text{m}5}_{i,t,h}$ in the following analysis due to the fact that the share of rounded histograms is particularly large in this case.

One explanation for the decision to round a forecast may be the amount of information that is available to all forecasters at the time a prediction is made rather than systematic differences between certain groups of panelists. In a fixed-event setting, the information set of a survey participant increases as h declines. In order to analyze the size of the groups of rounders and non-rounders, Table 5.2 summarizes the shares of non-rounded observations in the SPF data for each forecast horizon, that is,

$$S^{\mathcal{R}}_h = 100 \times \frac{\sum_i \sum_t D^{\mathcal{R}}_{i,t,h}}{\sum_i \sum_t D^{\mathcal{P}}_{i,t,h}}, \tag{5.14}$$

Figure 5.5: Share of rounded histograms (integer-based categorization)



Notes: The graphs depict the share of rounded histogram forecasts classified via the integer-based categorization from Eqn. (5.12), i.e., $\tilde{S}^{\mathrm{m}\tau} = 100 \times (\sum_i \sum_t \sum_h \tilde{D}^{\mathrm{m}\tau}_{i,t,h})/(\sum_i \sum_t \sum_h D^{\mathcal{P}}_{i,t,h})$ for $\tau \in \{1, \ldots, 10\}$, based on a pooled sample of observations across all forecast horizons for inflation (first row), output growth (second row) and unemployment (third row). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

where $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ denotes the preferred decimal- and integer-based rounding schemes described in Eqns. (5.7) and (5.12).

Table 5.2: Share of non-rounded observations

			Forecast horizon h							
SPF	Variable	Scheme	8	7	6	5	4	3	2	1
ECB	Inflation	D^{any}	13.65	13.15	13.76	12.30	12.42	12.22	12.70	12.29
		D^{m5}	22.93	22.43	22.48	21.76	22.53	22.78	22.67	22.48
	GDP growth	D^{any}	12.97	13.42	13.94	13.04	12.40	12.21	12.07	11.53
		D^{m5}	23.17	23.65	22.62	22.31	23.94	23.76	21.92	19.66
	Unemployment	D^{any}	12.15	12.43	12.45	11.74	12.31	12.20	13.29	12.16
		D^{m5}	22.22	21.52	21.82	23.48	21.75	21.49	22.37	20.05
FED	Inflation	D^{any}	4.07	3.92	4.38	4.03	4.07	4.73	5.07	4.76
		D^{m5}	21.36	22.71	21.21	19.65	17.92	20.39	21.33	21.14
	GDP growth	D^{any}	5.69	6.43	5.85	6.97	6.27	6.67	6.46	5.20
		D^{m5}	24.23	24.45	23.25	23.22	22.88	23.41	23.01	27.38
	Unemployment	D^{any}	10.09	8.22	8.56	8.30	9.05	7.83	9.09	6.42
		D^{m5}	22.37	22.37	26.13	24.02	26.70	23.96	28.23	28.88

Notes: For each forecast horizon, this table displays the share of non-rounded observations in the sample, i.e., $S^{\mathcal{R}}_h = 100\times(\sum_i \sum_t D^{\mathcal{R}}_{i,t,h})/(\sum_i \sum_t D^{\mathcal{P}}_{i,t,h})$ for the preferred decimal- and integer-based classification schemes $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ from Eqns. (5.7) and (5.12). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Table 5.2 shows that the share of non-rounded observations indicated by $D^{\text{any}}_{i,t,h}$ is relatively small in both surveys. Between 12–14% (ECB-SPF) and 4–10% (FED-SPF) of all histograms consist of probabilities that are stated with decimal numbers and are thus classified as being non-rounded. As will be discussed below, the larger value of S^{any}_h in the case of the ECB–SPF may be related to the reporting practices in both surveys. The share of non-rounded observations based on $D^{\text{m5}}_{i,t,h}$ is considerably larger and relatively similar in both versions of the SPF. In particular, the shares of non-rounded histograms in this case are 20–24% (ECB-SPF) and 18–29% (FED-SPF).[9] In both versions of the SPF, the share of non-rounders is relatively equal across outcome variables and forecast horizons. This suggests that the decision to round is not merely a result of more information being available as the target period approaches. The correlations between the outcomes of the distinct classification schemes are positive and statistically significant but vary across

[9] Naturally, if the probabilities are stated with decimal numbers, they cannot be multiples of an integer number. Conversely, if the probabilities are not multiples of a particular integer, they do not necessarily contain decimal numbers. Thus, the share of non-rounders isolated via $D^{\text{any}}_{i,t,h}$ is a subset of the share classified by $D^{\text{m5}}_{i,t,h}$.

both versions of the SPF, outcome variables and forecast horizons. A detailed account is omitted here for brevity.

Although the group of non-rounders is considerably smaller than the group of rounders in both surveys, the evidence documented in Table 5.2 shows that the share of non-rounders is relatively similar across outcome variables and forecast horizons. As a means to analyze the fluctuations in the status of active forecasters, Figure 5.6 depicts the time variation in the share of non-rounders for each variable across the predictions for both the current and the next year (defined analoguously to Eqn. (5.14)). As before, non-rounders are classified by means of either $D^{\text{any}}_{i,t,h}$ (first row) or $D^{\text{m5}}_{i,t,h}$ (second row).[10]

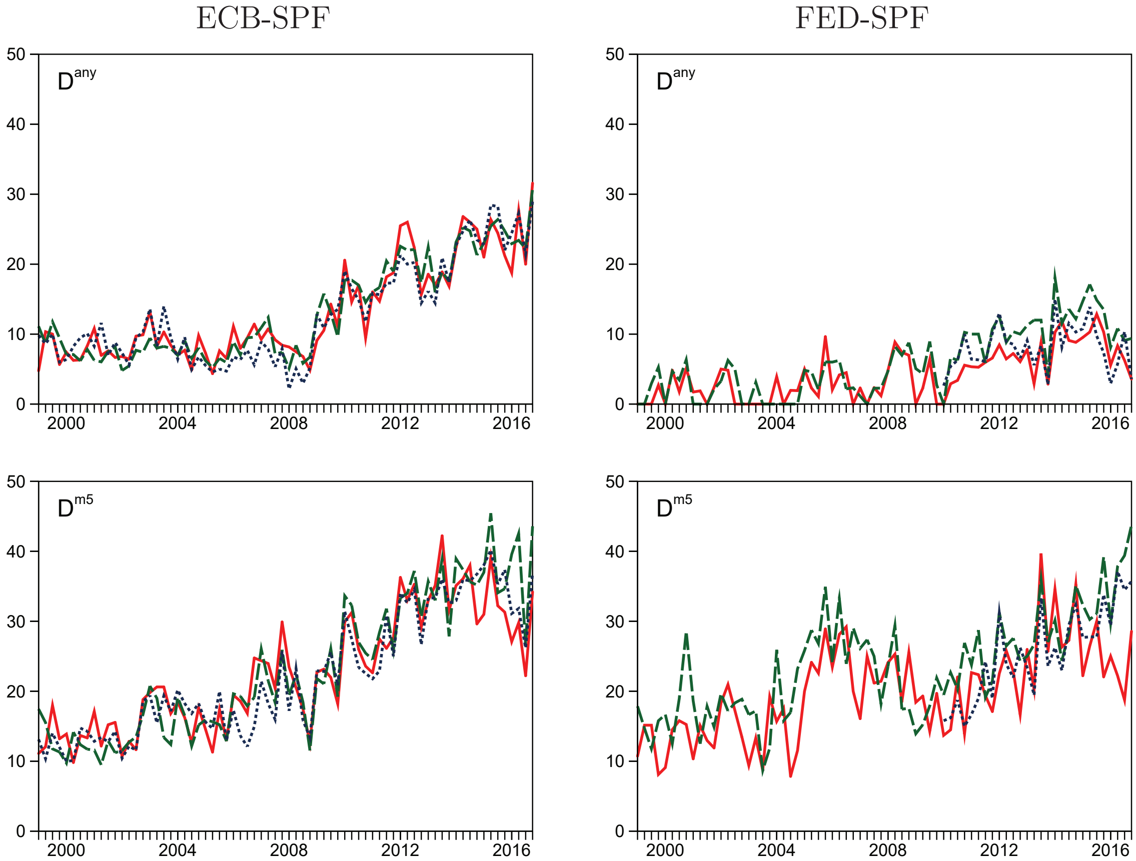
For each variable, the share of non-rounders in the ECB-SPF has considerably increased from approximately 5–15% of the cross-section during the initial years to 30–45% in recent survey periods. Over the same time period, the share of non-rounders in the FED-SPF has also increased, although it rarely exceeds 15% in the case of the categorization via $D^{\text{any}}_{i,t,h}$. In contrast, the share based on $D^{\text{m5}}_{i,t,h}$ is relatively similar in both versions of the SPF. This is in line with the previously documented evidence from Figures 5.2 and 5.5, which shows that participants of the FED-SPF rarely state probabilities in terms of decimal numbers, but more frequently not as multiples of five. Overall, Figure 5.6 documents an increase in the share of non-rounded histograms in both versions of the SPF during more recent years.

To investigate in which aspects the reported histograms of the non-rounders differ from those of the rounders, we examine the relation between rounding and two features of the histograms: First, we count the number of bins to which a forecaster assigns a nonzero probability, i.e., $K_{i,t,h}$. Second, we analyze the variance of the individual histograms, i.e., $\sigma^2_{i,t,h}$ from Eqn. (5.4). It has been hypothesized by Boero et al. (2015) that rounding may indicate inferior information or minor predictive ability. This should be reflected in the forecast performance of such individuals. Similarly, if the histograms of the non-rounders are more dispersed than those of the rounders, the variance misalignment should at least be partially explained by the rounding choices of the survey participants. As an example, consider the two histograms depicted in Figure 5.7.

The histogram depicted in the left plot corresponds to Example B from Figure 5.1. This forecast is unanimously classified as non-rounded based on $D^{\text{any}}_{i,t,h}$ and $D^{\text{m5}}_{i,t,h}$. Moreover, nonzero probabilities are assigned to each bin, such that $K_{i,t,h} = 12$. Based on Eqn. (5.4), the variance of this histogram is given by $\sigma^2_{i,t,h} = 0.72$. The right plot depicts what results if the probabilities in the reported histogram are artificially rounded to the nearest

[10]The participation and status of the individual survey participants is depicted in Figures 5.16 and 5.17 in the Appendix.

Figure 5.6: Time-variation in the share of non-rounders

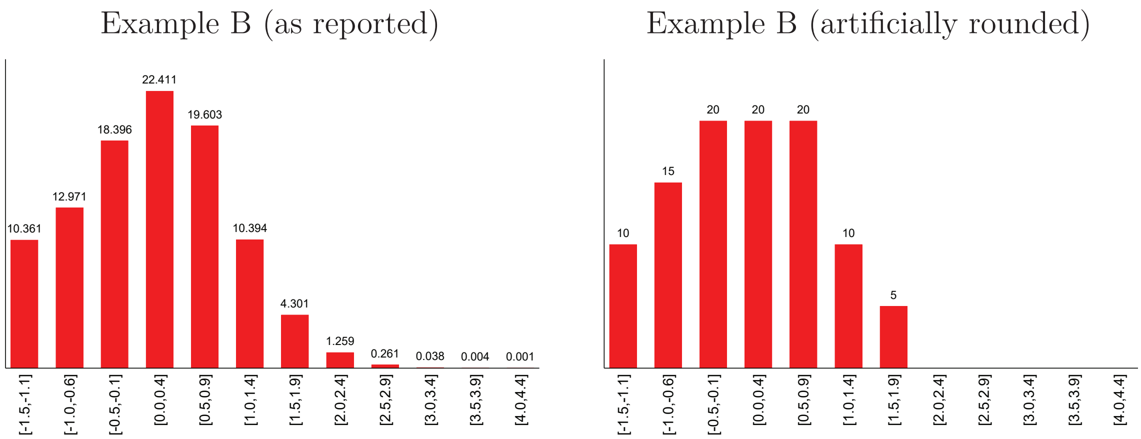


Notes: The graphs depict the share of non-rounded histogram forecasts for inflation (solid), output growth (dashed) and unemployment (dotted) based on $D^{\text{any}}_{i,t,h}$ (first row) and $D^{\text{m5}}_{i,t,h}$ (second row) for a pooled sample of observations across the predictions for the current ($h \leq 4$) and the next year ($h \geq 5$). The horizontal axis depicts the quarter during which predictions are reported. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

multiple of five. As a result, all of the probabilities in the right tail are rounded to zero, such that $K_{i,t,h}$ reduces to seven. In addition, the variance of the histogram reduces to $\sigma^2_{i,t,h} = 0.67$. This is a reduction by 7%. Thus, rounding has a considerable impact on the histogram width as measured by both $K_{i,t,h}$ and $\sigma^2_{i,t,h}$ in this particular case. To see if these effects also hold for the data set as a whole, a general analysis is provided below.[11]

[11] To disentangle the effect of rounding on the ex-ante variance from any other influence like the (unobserved) individual characteristics of the anonymous survey participants, we have conducted the artificial rounding exercise from Figure 5.7 for all non-rounders in the SPF data: For each histogram with $D^{\text{m5}}_{i,t,h} = 1$ we rounded the reported probabilities to multiples of five. After excluding observations where the artificially rounded probabilities do not sum to 100% we have found that the average variance from Eqn. (5.17) reduced by 7–10% (ECB-SPF) and 10–12% (FED-SPF) depending on the outcome variable. The average variance based on the artificially rounded histograms remained higher than the one of the rounders from

Figure 5.7: Artificial rounding exercise



Notes: The left plot depicts the one-quarter-ahead histogram forecast for the inflation rate reported by forecaster 102 in the 2016Q4 survey round of the ECB-SPF (see Figure 5.1). The right subfigure displays the result of artificially rounding the originally reported probabilities to multiples of five.

For both rounders and non-rounders, we calculate the average number of bins used by the individuals in each group,

$$\bar{K} = \frac{\sum_i \sum_t \sum_h K_{i,t,h} \times D^{\mathcal{R}}_{i,t,h}}{\sum_i \sum_t \sum_h D^{\mathcal{R}}_{i,t,h}} \tag{5.15}$$

and

$$\overline{\widetilde{K}} = \frac{\sum_i \sum_t \sum_h K_{i,t,h} \times \tilde{D}^{\mathcal{R}}_{i,t,h}}{\sum_i \sum_t \sum_h \tilde{D}^{\mathcal{R}}_{i,t,h}}, \tag{5.16}$$

where $\tilde{D}^{\mathcal{R}}_{i,t,h} = 1 - D^{\mathcal{R}}_{i,t,h}$. We also consider the average variances of both groups, i.e.,

$$\overline{\sigma^2} = \frac{\sum_i \sum_t \sum_h \sigma^2_{i,t,h} \times D^{\mathcal{R}}_{i,t,h}}{\sum_i \sum_t \sum_h D^{\mathcal{R}}_{i,t,h}} \tag{5.17}$$

and

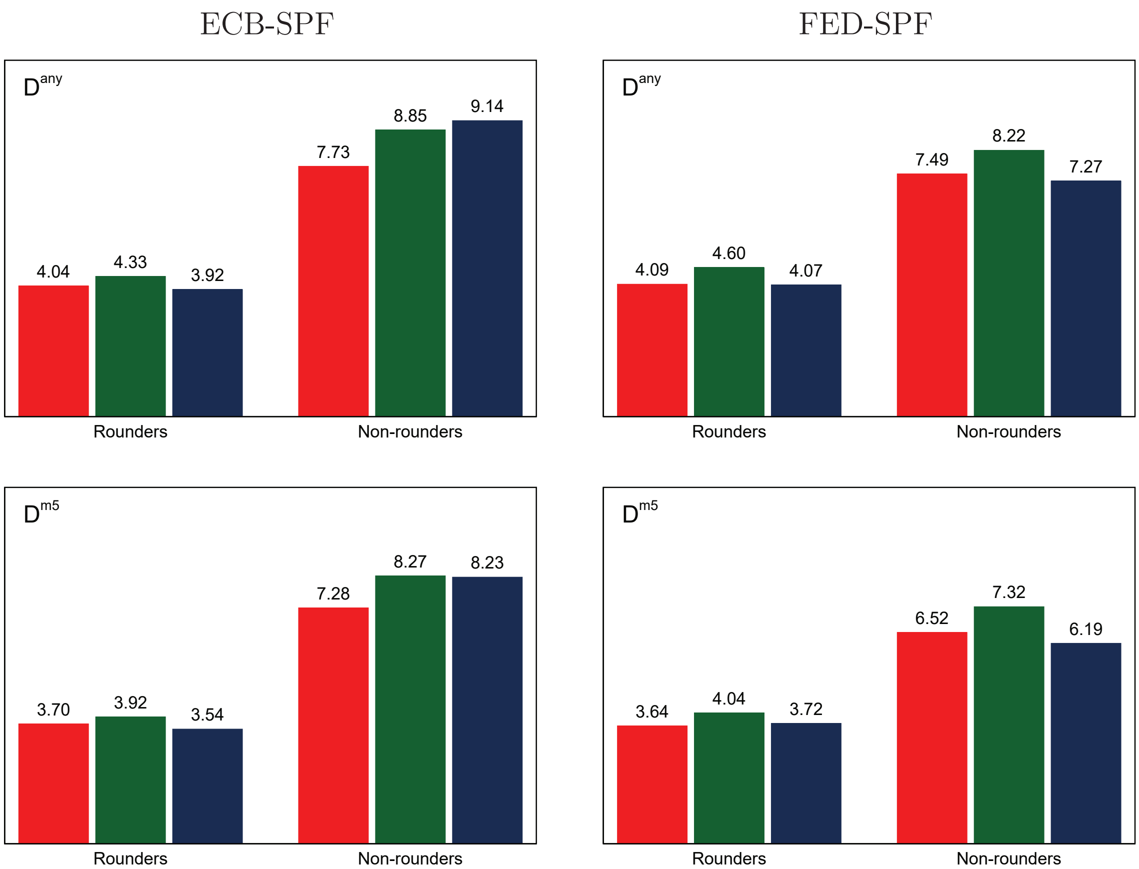
$$\overline{\widetilde{\sigma^2}} = \frac{\sum_i \sum_t \sum_h \sigma^2_{i,t,h} \times \tilde{D}^{\mathcal{R}}_{i,t,h}}{\sum_i \sum_t \sum_h \tilde{D}^{\mathcal{R}}_{i,t,h}}. \tag{5.18}$$

Note that it is unclear from an ex-ante point of view whether rounders or non-rounders report histograms with a higher dispersion. The results based on the decimal- and integer-

Eqn. (5.18), which suggests that other factors besides rounding explain part of the differences in the reported level of uncertainty. For brevity, these results are not reported in detail here.

based categorizations are depicted in Figures 5.8 and 5.9 for the average number of bins and variances, respectively.

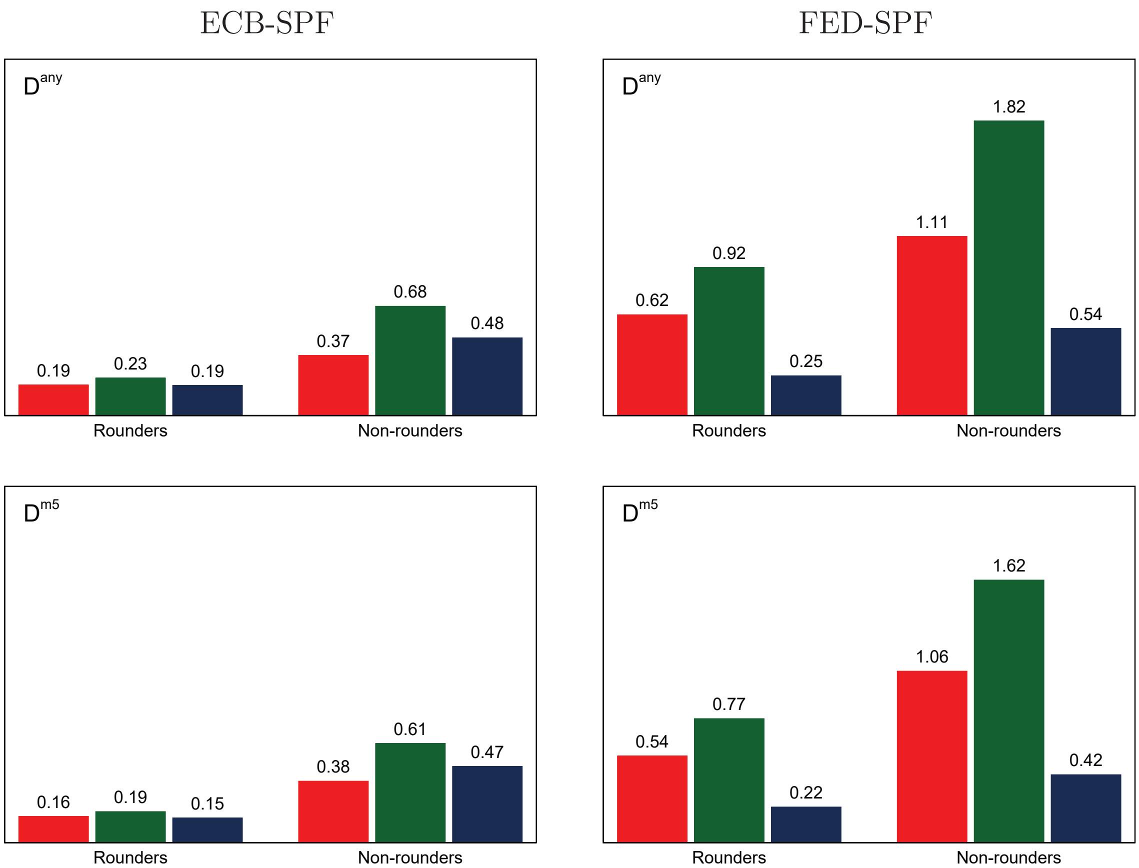
Figure 5.8: Average number of bins used by rounders and non-rounders



Notes: The graphs depict the average number of bins used by rounders and non-rounders based on the predictions for inflation, output growth and unemployment for a pooled sample of observations across forecasters, time periods and forecast horizons. Non-rounders are classified by means of $D^{\text{any}}_{i,t,h}$ (first row) or $D^{\text{m5}}_{i,t,h}$ (second row). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

As shown in Figure 5.8, the rounders in both the ECB- and FED-SPF assign nonzero probabilities to four bins on average, whereas the non-rounders use twice as many in most cases. This finding is remarkably robust across outcome variables and the employed categorization. The implication of this result is that rounding is related to the width of the reported histograms. Forecasters may round small probabilities in the tails of the histogram to zero as illustrated in Figure 5.7. Similarly, the variances of the non-rounders are, on average, approximately twice as large as those of the rounders. However, there is substantial heterogeneity in the level of ex-ante uncertainty across surveys and outcome variables. This squares with the evidence from Figure 5.4. Our results suggest that the

Figure 5.9: Average ex-ante variances reported by rounders and non-rounders



Notes: The graphs depict the average across the ex-ante variances reported by rounders and non-rounders based on the predictions for inflation, output growth and unemployment for a pooled sample of observations across forecasters, time periods and forecast horizons. Non-rounders are classified by means of $D^{\text{any}}_{i,t,h}$ (first row) or $D^{\text{m5}}_{i,t,h}$ (second row). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

degree of the variance misalignment may be related to the rounding behavior of individual forecasters. Since K, the number of available bins, is considerably larger in the ECB-SPF across all variables, the similarities in the results across both versions of the SPF suggest that these findings are not just an inherent consequence of the different survey designs. We examine these issues in detail in Section 5.5.3.

5.5.2 Analysis of Variance Misalignment

The evidence reported in the previous subsection suggests a relationship between the dispersion of the reported histograms and the rounding choices of individual forecasters. Based on this observation, we compare the ex-ante and ex-post variances of the SPF participants while accounting for the fact that there may be differences in the degree of the variance misalignment between rounders and non-rounders. Moreover, it could be the case that the rounding choices of individual forecasters are related to the information available to them. Ruud et al. (2014) show that rounding is more likely to occur in situations in which individuals have noisy information about the outcome. In the case of survey-based fixed-event forecasts, the survey participants should become better informed as the forecast horizon shrinks during successive survey rounds. If this is the case, the differences in the variance misalignment may be related to the forecast horizon. This hypothesis is examined next. We measure the ex-ante uncertainty of forecaster i at forecast horizon h by means of the individual-specific average variance, which is defined as

$$\overline{\sigma^2_{i,h}} = \frac{1}{T_{i,h}} \sum_{t=1}^{T_{i,h}} \sigma^2_{i,t,h}, \tag{5.19}$$

where $T_{i,h} = \sum_{t=1}^{T} D^{\mathcal{P}}_{i,t,h}$ indicates the number of times forecaster i has reported h-step-ahead predictions and $\sigma^2_{i,t,h}$ denotes the variance from Eqn. (5.4). In order to analyze the degree of the variance misalignment in the SPF, the ex-ante uncertainty from Eqn. (5.19) is compared to the mean squared error (MSE), as given by

$$\text{MSE}_{i,h} = \frac{1}{T_{i,h}} \sum_{t=1}^{T_{i,h}} e^2_{i,t,h}. \tag{5.20}$$

The MSE in Eqn. (5.20) is based on the individual forecast errors,

$$e_{i,t,h} = x_t - \mu_{i,t,h}, \tag{5.21}$$

with x_t denoting the realization of the outcome variable and $\mu_{i,t,h}$ indicating the mean of forecaster i's histogram as defined in Eqn. (5.2). To compare the ex-post and ex-ante variances across all survey participants, we compute the average misalignment ratio,

$$m_h = \frac{1}{N_h} \sum_{i=1}^{N_h} \frac{\text{MSE}_{i,h}}{\overline{\sigma^2_{i,h}}}, \tag{5.22}$$

for each forecast horizon, where N_h denotes the number of survey participants who report h-step-ahead histogram forecasts for outcome variable x_t. If forecasters provide an accurate ex-ante quantification of the average size of their forecast errors, the value of the statistic in Eqn. (5.22) equals unity.[12] Values above unity are typically interpreted as

[12]Note that the statistic in Eqn. (5.22) differs from the one employed in Clements (2014) where the root MSE and the standard deviations are used to compute a similar ratio. Due to the nonlinearity of this

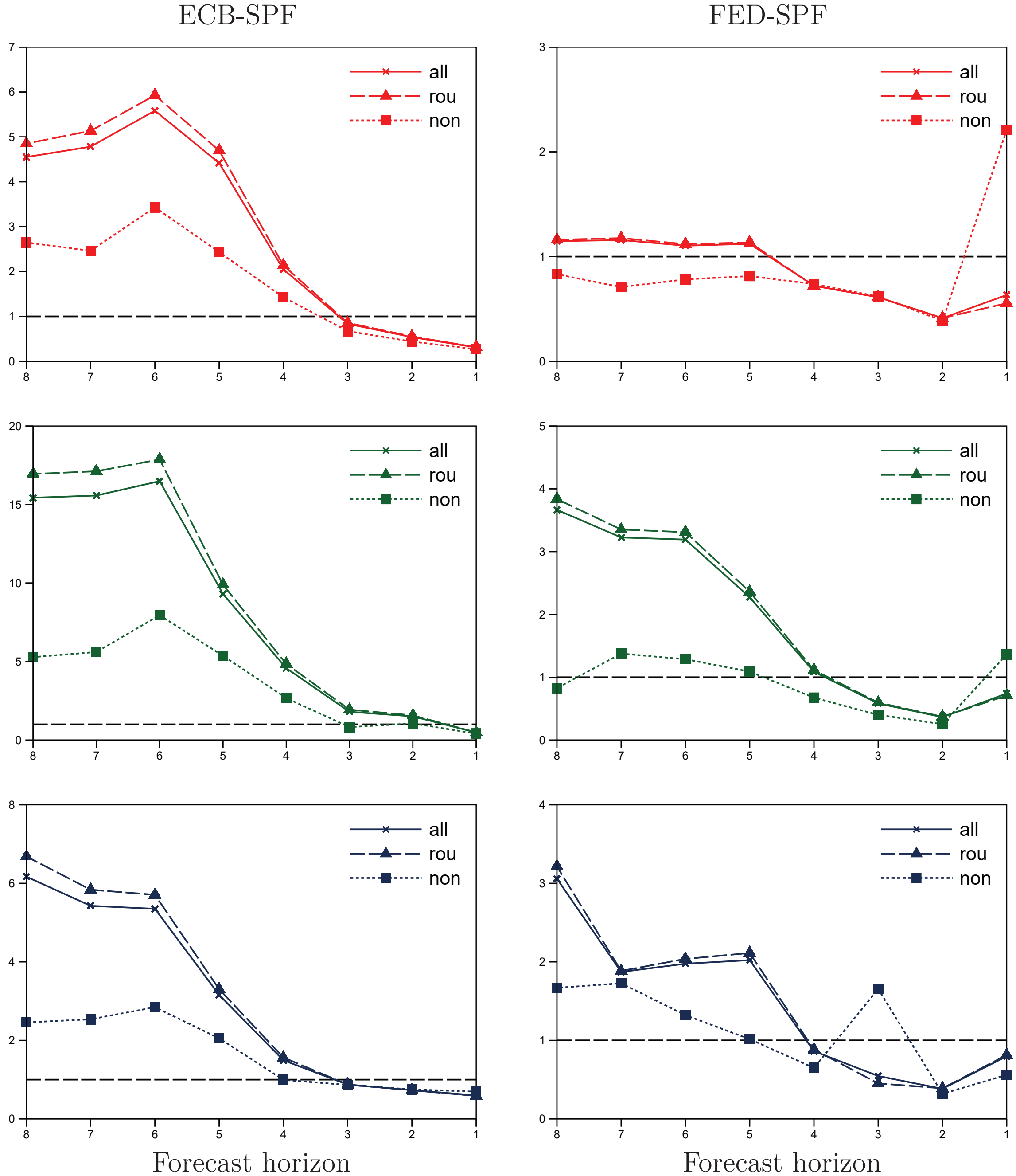
evidence of 'overconfidence', i.e., cases where the ex-ante uncertainty is, on average, too small compared to the ex-post uncertainty. We compute the average misalignment ratio across all forecasters as well as separate ratios for the rounders and non-rounders based on both $D^{\text{any}}_{i,t,h}$ and $D^{\text{m5}}_{i,t,h}$. The results are summarized in Figures 5.10 and 5.11, respectively.

The evidence for the entire cross-section show that the variance misalignment can be diagnosed in both versions of the SPF. The values of the m_h-ratio in both surveys tend to be substantially larger than unity at forecast horizons of one year or more, i.e., the ex-post and ex-ante variances are better aligned as the target period approaches. In particular, there is a notable drop in m_h as the forecast horizon diminishes from five to four quarters ahead. As discussed in Lahiri and Sheng (2008), this may be related to the availability of first releases of data for x_t for the respective year or alternative sources of information about the outcome. At the shortest forecast horizons, the ex-ante variances are frequently larger than the ex-post variances. In these cases, forecasters overstate their ex-ante uncertainty compared to the forecast errors and should, on average, reduce the variance of their histogram close to the target. These findings square with similar evidence documented in Giordani and Söderlind (2003, 2006) and Clements (2014, 2016) for the FED-SPF as well as Kenny et al. (2014) and Krüger (2017) for the ECB-SPF. In particular, we confirm the result of Clements (2014) that forecasters in the SPF report higher ex-ante than ex-post variances at short forecast horizons. The observed pattern is remarkably consistent across variables. The inflation rate forecasts in the FED-SPF are an exception since they are relatively well aligned even at long forecast horizons. Moreover, in most cases the degree of the variance misalignment is larger in the ECB-SPF than in the FED-SPF.

Empirical studies on the variance misalignment in surveys of macroeconomic expectations evaluate the entire cross-section of forecasters. By isolating rounders and non-rounders in the SPF by means of the $D^{\text{any}}_{i,t,h}$ or $D^{\text{m5}}_{i,t,h}$ categorizations, we find that the average ratio of the non-rounders is much closer to unity at forecast horizons of one year or more, which are particularly those horizons for which the studies listed above tend to find the most substantial evidence of 'overconfidence'. In contrast, the average ratios of the rounders and non-rounders are relatively similar as the target period approaches. In sum, the results indicate that the ex-ante and ex-post variances of the non-rounders are better aligned than those of the rounders at forecast horizons of one year or more. Thus, it appears the variance misalignment is at least partially explained by the rounding choices of the SPF participants. Rounding may affect both the numerator and the denominator of the statistic in Eqn. (5.22). On the one hand, the histogram mean can be affected. On

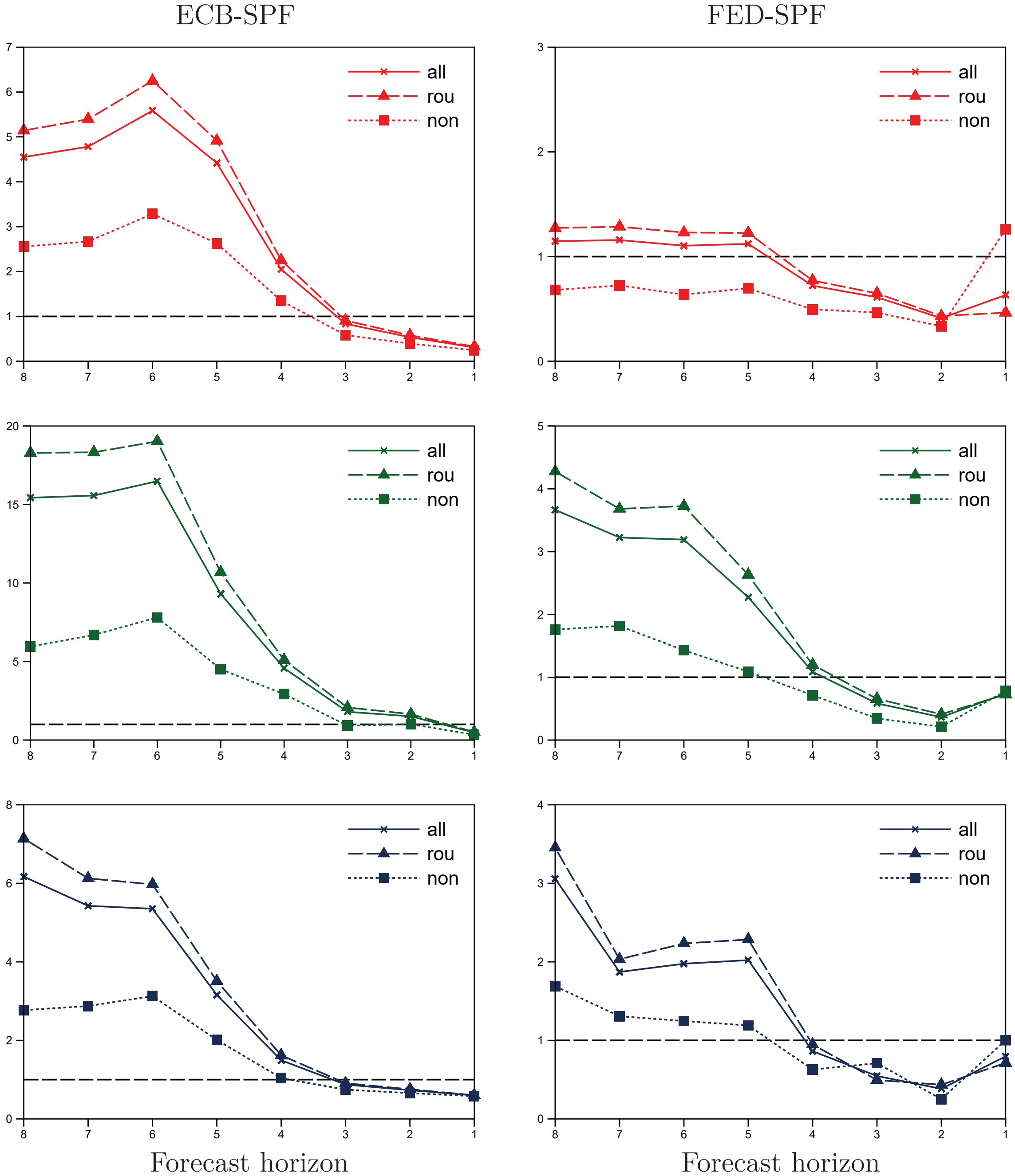
transformation, the two statistics cannot be directly compared. Lahiri et al. (2015) discuss the distinct interpretations that arise due to the ordering by means of which aggregation and the root-transformation are applied. To avoid this type of ambiguity, we opt for employing the variance and the MSE instead.

Figure 5.10: Variance misalignment in the SPF data (decimal-based categorization)



Notes: Each plot depicts the misalignment ratio m_h from Eqn. (5.22) for inflation (first row), output growth (second row) and unemployment (third row) in the ECB- (first column) and FED-SPF (second column). In addition to the average ratio for the entire cross section (solid line), each plot depicts separate ratios for rounders (dashed line) and non-rounders (dotted line). Non-rounders are classified by means of $D_{i,t,h}^{\text{any}}$. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.11: Variance misalignment in the SPF data (integer-based categorization)



Notes: Each plot depicts the misalignment ratio m_h from Eqn. (5.22) for inflation (first row), output growth (second row) and unemployment (third row) in the ECB- (first column) and FED-SPF (second column). In addition to the average ratio for the entire cross section (solid line), each plot depicts separate ratios for rounders (dashed line) and non-rounders (dotted line). Non-rounders are classified by means of $D^{\mathrm{m5}}_{i,t,h}$. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

the other hand, rounding may be related to the ex-ante uncertainty as measured by the variance of the histogram. This is analyzed in the next subsection.

5.5.3 Differences in Histogram Characteristics

The improved alignment of the ex-ante and ex-post variances of the non-rounders documented in the previous subsection may be both due to a higher dispersion of the reported histograms or be the consequence of smaller forecast errors. The evidence from Figures 5.8 and 5.9 shows that the subjective uncertainties reported by the non-rounders are considerably larger, which means that the denominator of the ratio in Eqn. (5.22) is larger for this particular group. To shed light on the potential reasons for the misalignment of variances, we analyze the forecast performance and histogram characteristics of rounders and non-rounders below. For evaluating the impact of (non-)rounding, we estimate horizon-specific regressions of the form

$$y_{i,t,h} = \alpha_h + \beta_h D^{\mathcal{R}}_{i,t,h} + \gamma_{2,h} D2_{t,h} + \ldots + \gamma_{T,h} DT_{t,h} + \varepsilon_{i,t,h}, \tag{5.23}$$

where $y_{i,t,h} \in \{K_{i,t,h}, \sigma^2_{i,t,h}, |e_{i,t,h}|, e^2_{i,t,h}\}$ denotes distinct histogram characteristics, variation measures and loss functions, respectively, $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ indicates the employed categorization for (non-)rounding and $\varepsilon_{i,t,h}$ is the error term. The first group of histogram characteristics consists of variables that capture the histogram width, i.e., the number of bins used by forecasters, $K_{i,t,h}$, and the individual variance defined in Eqn. (5.4). These variables are observable ex-ante and affect the denominator of Eqn. (5.22). The second group captures the individual ex-post forecast performance based on the realizations and the histogram means. In particular, we consider the absolute individual forecast errors, $|e_{i,t,h}| = |x_t - \mu_{i,t,h}|$, and the squared forecast errors, $e^2_{i,t,h} = (x_t - \mu_{i,t,h})^2$. Both are related to the numerator of the ratio in Eqn. (5.22). We include time-fixed effects $D2_{t,h}, \ldots, DT_{t,h}$ in order to capture unobserved sources of heterogeneity that vary across time, but affect the survey participants in the same way. In particular, this includes changes in the design of the survey questionnaire.

In Eqn. (5.23), each candidate variable for $y_{i,t,h}$ is regressed on $D^{\mathcal{R}}_{i,t,h}$, i.e., the indicator for non-rounding. The slope coefficients $\beta_8, \ldots, \beta_1$ capture the differences in the histogram characteristics of non-rounders and rounders for distinct forecast horizons $h \in \{8, 7, \ldots, 1\}$. The parameter vector $(\alpha_h, \beta_h, \gamma_{2,h}, \ldots, \gamma_{T,h})'$ is estimated via ordinary least squares (OLS). The sample size used in the estimation for each h is reported in Table 5.1. Since the data used in each regression are observed at the annual frequency, the error terms in Eqn. (5.23) are correlated across time periods due to the overlapping forecast horizons in cases where $h > 4$. In order to account for the autocorrelation patterns in the data, we apply the variance-covariance estimator by Newey and West (1987).

Figures 5.12–5.15 display the estimates of β_h over h for each outcome variable. The significant and insignificant estimates are highlighted differently. In particular, a diamond '$\diamond$' indicates that the estimate is significantly different from zero at the 5% critical level against a two-sided alternative. Generally, the results are robust to the choice of the classification scheme.[13] Note that the estimates for the FED-SPF are more strongly affected by individual observations due to the smaller share of non-rounders in this survey (see Table 5.2).

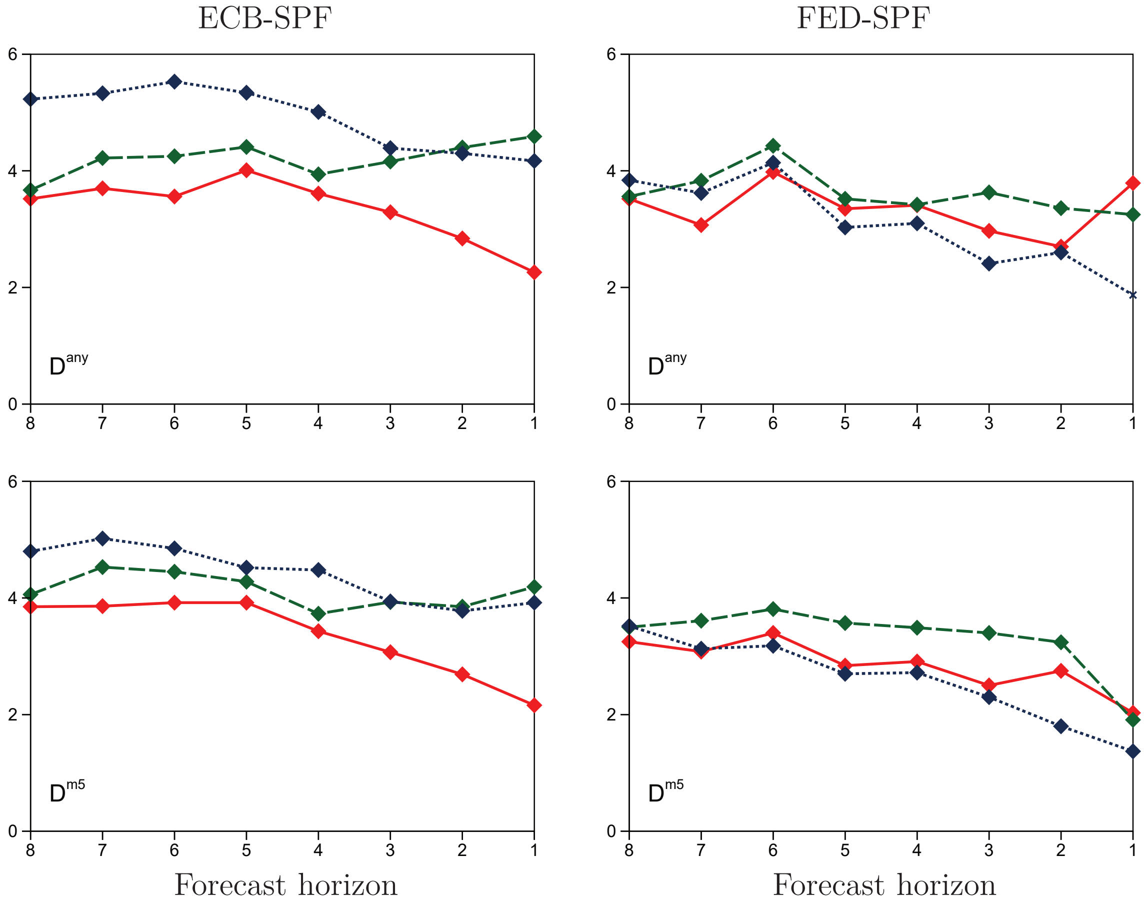
Differences in Individual Histogram Ranges and Variances

The evidence from Figures 5.8 and 5.9 shows that the non-rounders use more of the available bins and report higher variances than the rounders for a pooled sample of observations based on all forecast horizons. Yet it is not clear whether these differences vary with h. This may be the case if rounders and non-rounders update their information sets at different frequencies, e.g., due to heterogeneity in the level of information stickiness or differences in the horizons forecasters are concerned with as part of their principal occupation. In order to analyze the importance of the forecast horizon, Figures 5.12 and 5.13 depict the estimates of β_h that result when either the employed number of bins, $K_{i,t,h}$, or the individual ex-ante variance, $\sigma^2_{i,t,h}$, are used as the dependent variable in the model from Eqn. (5.23). Forecasters are classified as non-rounders based on either $D^{\text{any}}_{i,t,h}$ (first row) or $D^{\text{m5}}_{i,t,h}$ (second row).

The results for $K_{i,t,h}$ confirm the evidence from Figure 5.8. For each forecast horizon, the non-rounders in both surveys use significantly more bins than the rounders. The finding that non-rounders fill in a larger number of bins is also found for all particular forecast horizons. On average, the difference is approximately equal to four bins. However, in most cases the differences become less pronounced as the forecast horizon diminishes. Thus, the larger variances of the non-rounders are revised downwards more strongly as the target is approached during the forecasting process. This pattern is particularly apparent for the estimates based on the inflation and unemployment rate forecasts in the ECB-SPF. The values of the adjusted R^2-statistics (not shown) are lower in the FED-SPF than in the ECB-SPF. In the former case, the models explain 8–43% of the variation in $K_{i,t,h}$, whereas 19–57% are explained in the latter case. It could be that differences in the survey methodology are the reason for the improved goodness of fit. The reporting practices permitted in the case of the ECB-SPF may allow the employed rounding classification

[13] The results for the other categorizations are reported in Figures 5.18-5.21 in the Appendix.

Figure 5.12: Deviations in the number of bins used by non-rounders and rounders

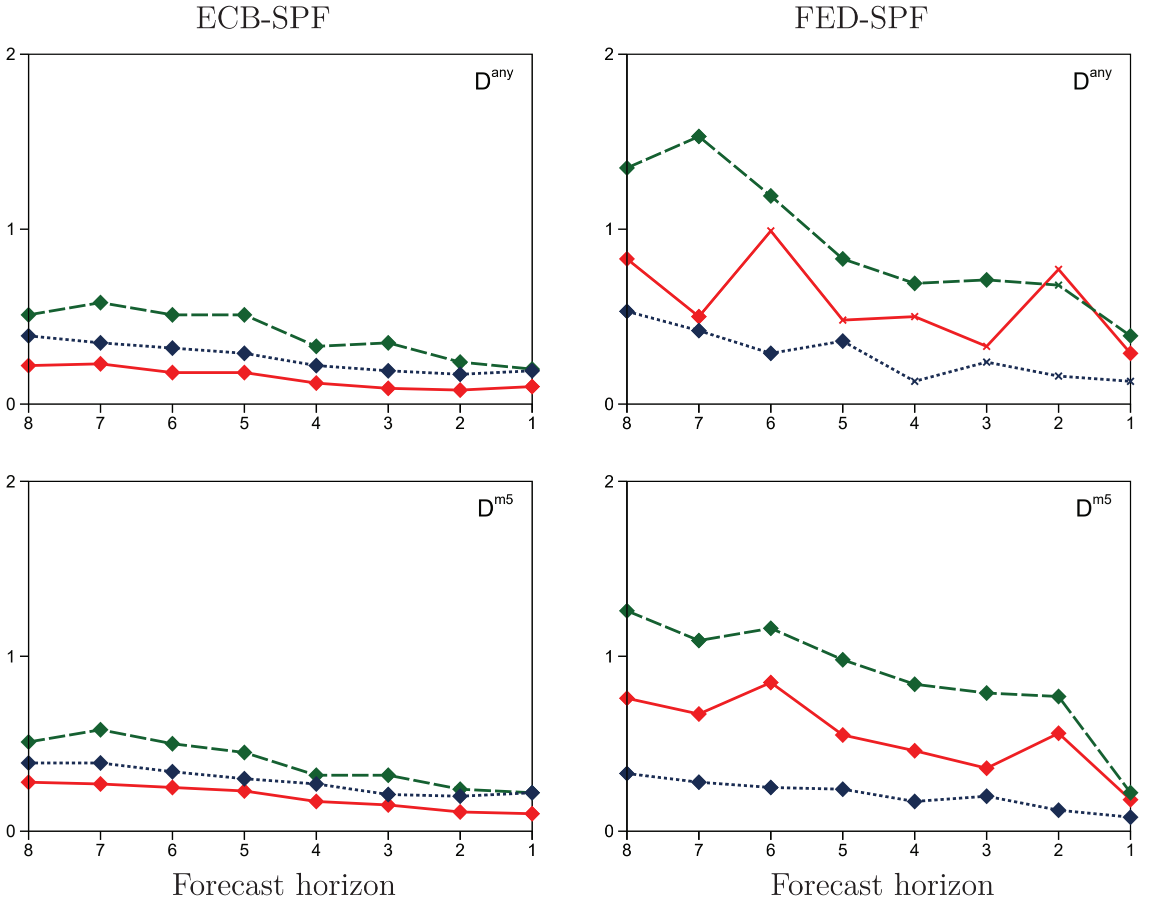


Notes: For each forecast horizon, the graphs depict the difference in the number of bins used by non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $K_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '⋄' indicates that the number of bins used is distinct among non-rounders and rounders. The significance level is 5%. A cross '×' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ denotes either the preferred decimal- (first row) or integer-based (second row) categorization from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

to isolate uninformed from informed survey participants. In contrast, the categorizations might be less precise in the case of the FED-SPF due to the fact that the panelists are required to fill in the questionnaire manually.

Overall, the results from Figure 5.12 suggest that the non-rounders use significantly more bins than the rounders and that this difference frequently becomes smaller as the horizon is dimnishing. Moreover, the evidence suggests that our results are not strongly affected by the choice of the employed classification scheme.

Figure 5.13: Deviations in the variances reported by non-rounders and rounders



Notes: For each forecast horizon, the graphs depict the difference in the ex-ante variances reported by non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $\sigma^2_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the reported variance forecasts are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ denotes either the preferred decimal- (first row) or integer-based (second row) categorization from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

The evidence that is obtained when $\sigma^2_{i,t,h}$ is used as the dependent variable is in line with Figure 5.9 in the sense that non-rounders report significantly wider histograms. For the decimal-based categorization, the decreasing pattern of the estimated slope coefficients from Figure 5.8 is visible here as well. The goodness of fit is typically smaller than in the case of $K_{i,t,h}$ (1–22% and 7–42% for the FED- and ECB-SPF, respectively). In both surveys, the differences in the variances tend to decline in both magnitude and significance as the target approaches. The estimates for the FED-SPF are driven by a smaller number of individual observations than in the case of the ECB-SPF.

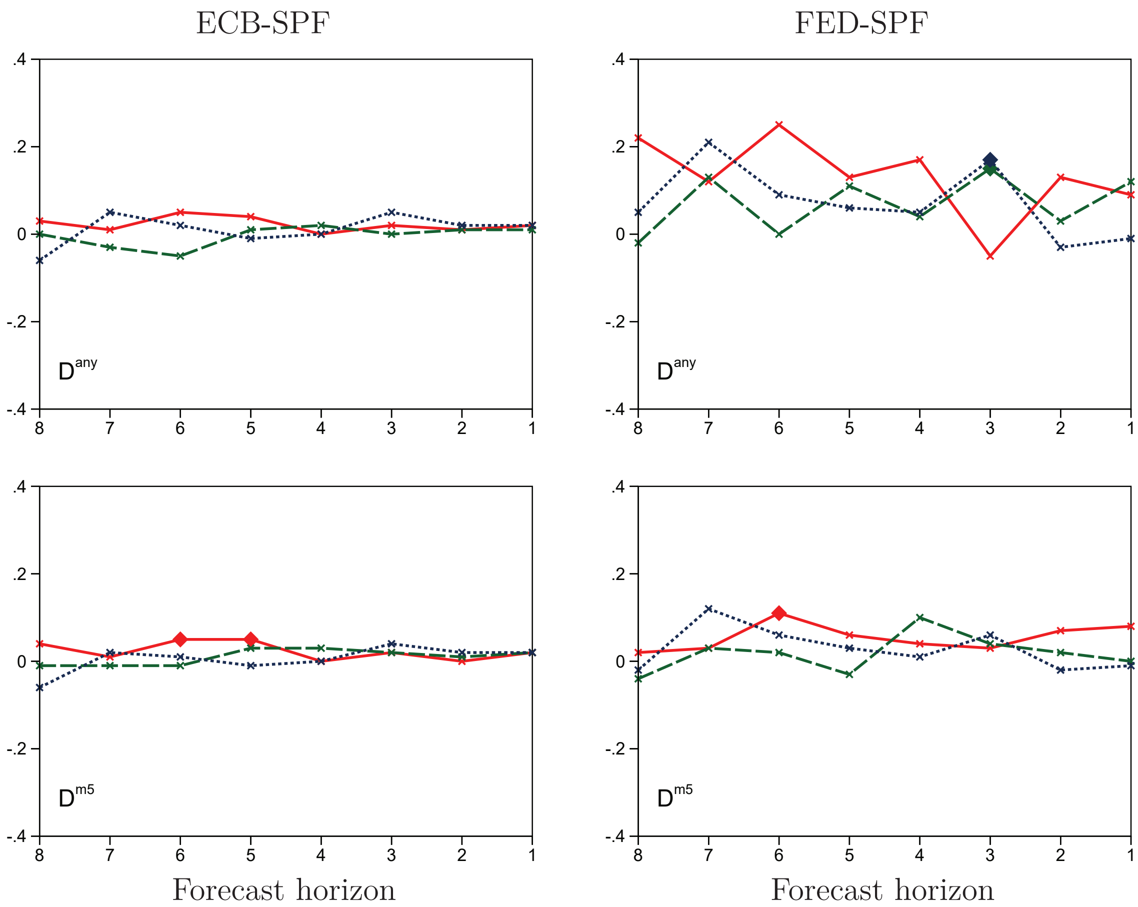
In sum, the results for $K_{i,t,h}$ and $\sigma^2_{i,t,h}$ confirm that non-rounders in the SPF use more bins and report larger variance forecasts than the rounders. In most cases, this implies that the denominator of the m_h-statistic from Eqn. (5.22) is larger for the non-rounders. The differences become smaller as the target approaches, which provides a potential explanation for the similar alignment of the ex-post and ex-ante variances reported by rounders and non-rounders at the short forecast horizons (see Figures 5.10 and 5.11). The results are robust to the choice of the categorization if the share of non-rounders in the cross-section is sufficiently large. This can be observed for both surveys, but it is particularly visible in the ECB-SPF, which contains a larger number of individuals that we classify as non-rounders than the FED-SPF.

Differences in Forecast Errors

The results from Figures 5.12 and 5.13 reveal that the histograms reported by the non-rounders are more dispersed than those of the rounders. This is particularly the case for forecast horizons of one year or more. These horizons correspond to those for which the difference in the variance misalignment between both groups is particularly large (see Figures 5.10 and 5.11). Apart from the denominator of Eqn. (5.22), the numerator can also be the reason for the variance misalignment. The size of the numerator depends on the individual forecast errors. To analyze the impact of (non-)rounding on the predictive accuracy of the histogram means, Figures 5.14 and 5.15 depict the estimated slope coefficients when either absolute or squared forecast errors, $|e_{i,t,h}|$ and $e^2_{i,t,h}$, are considered as the dependent variable in Eqn. (5.23).

We find no significant differences in the ex-post forecast performance of rounders and non-rounders in terms of either absolute or squared forecast errors. In the case of the ECB-SPF, the estimates of β_h are very close to zero for both types of prediction errors and both categorizations. The results for the FED-SPF are more erratic. Nonetheless, the null hypothesis that β_h equals zero is not rejected in almost all cases. Overall, the results suggest that the histogram mean is robust to the rounding choices of the survey participants. This is in line with the evidence of Engelberg et al. (2009), who show that rounding has little impact on the mean of a forecaster's subjective distribution. Similarly, Binder (2017) decomposes disagreement, defined as the cross-sectional dispersion of the point forecasts, into the contributions of the rounding and non-rounding group and shows that almost all of the cross-sectional variability can be ascribed to variation within the respective groups, i.e., rounders and non-rounders, meaning that the group-specific means do not differ substantially.

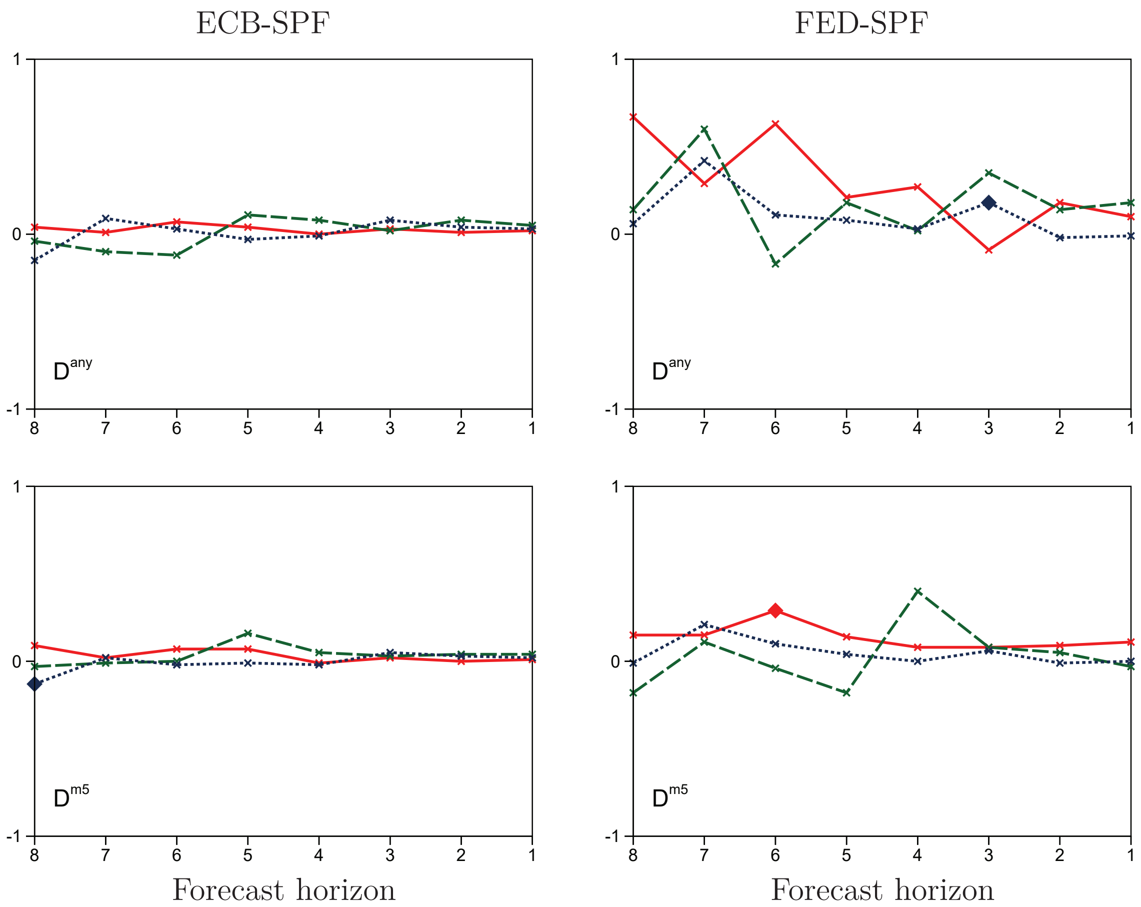
Figure 5.14: Deviations in the absolute forecast errors of non-rounders and rounders



Notes: For each forecast horizon, the graphs depict the difference in the ex-post absolute forecast errors of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $|e_{i,t,h}|$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the absolute prediction errors are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ denotes either the preferred decimal- (first row) or integer-based (second row) categorization from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

To summarize, our results suggest that the ex-ante and ex-post variances of SPF participants deviate substantially at forecast horizons of one year or more. This misalignment can be at least partially explained by the rounding choices of the panelists. In particular, we show that the variances of the non-rounders are better aligned due to the fact that this group of forecasters reports larger ex-ante variances but does not substantially differ from the rounders in terms of ex-post prediction errors. Thus, rounding choices

Figure 5.15: Deviations in the squared forecast errors of non-rounders and rounders



Notes: For each forecast horizon, the graphs depict the difference in the ex-post squared forecast errors of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $e^2_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the squared prediction errors are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{m5}}_{i,t,h}\}$ denotes either the preferred decimal- (first row) or integer-based (second row) categorization from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

affect the denominator of the misalignment ratio in Eqn. (5.22), but not the numerator.[14] The implication of this finding is that a better calibrated quantification of ex-ante uncertainty can be obtained by discarding strongly rounded histograms and focusing on the

[14]We have also analyzed whether the *degree* of (non-)rounding contains information about certain histogram characteristics by replacing $D^{\mathcal{R}}_{i,t,h}$ in Eqn. (5.23) with the average number of decimals per histogram forecast. The results are remarkably similar to our main results, i.e., each additional decimal numbers is associated with a significantly wider histogram in terms of both the number of bins and the ex-ante variance. In contrast, the average number of decimals has no predictive power for either absolute or squared forecast errors.

non-rounders. The share of non-rounded responses has been increasing recently as seen in Figure 5.6.

In additional regressions that are reported in Figures 5.22 and 5.23 in the Appendix, we analyze whether rounders and non-rounders differ in terms of the forecast performance of the entire histogram as measured by the quadratic probability score (QPS) and the ranked probability score (RPS) as discussed in Boero et al. (2011). The evidence suggests that the histograms of the non-rounders tend to outperform those of the rounders at long forecast horizons. However, the results vary both across versions of the SPF and outcome variables. We do not focus on these findings because they are not directly related to the analysis of the variance misalignment.[15]

5.6 Discussion

The previously documented findings suggest that to achieve a higher predictive accuracy of the ex-ante variances, it might pay if survey designers increase the number of non-rounders among the participants. Moreover, the large difference in the share of non-rounders in the ECB- and FED-SPF based on the decimal-based categorization, i.e., D^{any} (see Table 5.2) suggests that reporting techniques may play a role in the decision of a forecaster to report rounded numbers. Participants in the ECB-SPF can process and submit their responses online, whereas forecasters in the FED-SPF are required to print out the questionnaire and report their forecasts in a hand-written form. This may appear tedious to some non-rounders and induce them to report rounded probabilities instead. If this is the case, surveys of macroeconomic expectations should be designed in such a way that its participants can submit their forecasts with as little effort as possible. Nonetheless, the ECB-SPF sample contains the responses of a considerable number of rounders. This suggests that additional factors such as information deficiencies or ambiguity may play a role.

5.6.1 Rounding Versus Judgment

It may be the case that rounding choices reflect the fact that some survey participants use formal models to arrive at their forecasts, whereas others rely more on judgment and intuition. In order to shed light on the reporting practices of its participants, the

[15] In unreported regressions we have also considered higher moments and found no clear evidence for substantial deviations in the skewness of the histograms reported by both groups. On average, the SPF histograms tend to be relatively symmetric. Following Andrade et al. (2015), we have found that the histograms of the non-rounders exhibit a higher kurtosis than those of the rounders.

ECB-SPF conducted two special surveys in 2008 and 2013. Among other questions, respondents were asked if their probabilities are based on a model, judgment or a mixture of the two. In the first special survey, 79% of the survey participants answered that their reported probabilities are judgment-based, whereas the remaining panelists replied that they are derived from a formal model or a functional form (ECB, 2009). Interestingly, the fraction of forecasters who stated that they rely entirely on judgment is very close to the relative frequency of rounded observations classified by means of D^{m5} (see Table 5.2). In the second survey, the share of forecasters who indicated that their reported probabilities are based on judgment varies between 68% for the medium-term inflation and GDP growth forecasts and 79% for the short-term unemployment rate forecasts (ECB, 2014b). On average, the predominance of forecasters who rely on judgment has slightly declined compared to the first special survey. This squares with the increase in the share of non-rounders in recent survey periods depicted in Figure 5.6.[16] Notably, the share of forecasters who replied that they compute their probabilites only for the SPF (79%), as opposed to producing them for purposes related to their regular work, is the same as the fraction of forecasters who stated that they rely on judgment. Consequently, it is also very similar to the share of rounders as measured by $\tilde{S}^{\mathrm{m5}}$.

It is tempting to examine the link between the responses in the special surveys and the rounding choices in the quarterly SPF questionnaires. The questions in these surveys refer to the fixed-horizon forecasts, i.e., predictions with a constant forecast horizon. Thus, we consider the share of non-rounders for the fixed-horizon forecasts reported in the surveys that correspond to the dates of the special surveys, i.e., 2009Q4 and 2013Q3. Note, however, that the forecast horizons do not exactly match those from the special surveys. Moreover, the number of forecasters in the 2013Q3 survey is not identical to the number of responses from the second special survey. Nonetheless, the share of non-rounders in 2009Q4 based on D^{any} (19–26% depending on the variable and horizon) closely mirrors the share of forecasters who reported that they use some sort of model when they report their probabilities (21%).[17] The share of non-rounders classified by means of D^{any} in 2013Q3 (19–31%) is relatively similar to the fraction of forecasters who replied that they use either a model or a combination of model and judgment in the second special survey (26–33%).

[16]The share of cases where judgment is applied is considerably smaller for the point predictions and rarely exceeds 50% in the first special survey. In the second special survey, the fraction of point predictions based on judgment has further declined. In particular, the share of forecasters who replied that their point forecasts are essentially judgment-based is 35% or less for the forecast horizons of at most three years ahead. The share is considerably larger for the long-term predictions, but remains below 50%. Out of the remaining panelists, 14–28% indicated that their point predictions are model-based, while the remaining 25–60% replied that they use a mixture of judgment and models.

[17]We consider both the category 'econometric model' and what is referred by the ECB as a 'functional form' as cases where forecasters employ some generic form of model.

Thus, it appears that there is a close association between our distinction of rounders and non-rounders on the one hand and the non-judgment versus judgment-based forecast grouping documented in the special survey of the ECB on the other hand.[18] However, it is possible that this is entirely a matter of coincidence. Forecasters whom we classify as being non-rounders may be entirely different from those who report that they use formal models.[19]

5.6.2 Expert Versus Consumer Surveys

Using an approach similar to our integer-based categorization, Binder (2017) analyzes the relationship between histogram width and rounding in the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. She finds that the average histogram width of the rounders is approximately twice as large as that of the non-rounders. In contrast, we find that the histograms of the *non-rounders* are more dispersed. However, as will be discussed below, there are important distinctions between both analyses. Moreover, we show that our categorizations and the one used by Binder (2017) isolate distinct groups of survey participants.

First, we consider professional forecasters, whereas Binder (2017) focuses on consumers. There may be systematic differences in the way that each group computes their predictions. As discussed in the previous subsection, survey participants may rely on either formal models or judgment in the forecasting process. It seems likely that the relative importance of judgmental forecasting is higher for consumers than it is for experts. Second, we classify the SPF participants as rounders or non-rounders based on their histogram forecasts. Binder (2017) focuses on the point forecasts instead. For consumer surveys, this may be advantageous since consumers who are not expert forecasters may focus their attention on approximating the first moment and put less effort into a sophisticated quantification of higher moments. The categorizations employed in our study have the advantage that they are based on more than just one number due to the fact that almost all SPF participants assign nonzero probabilities to multiple bins. Thus, the two approaches can be considered as complementary to each other. However, it is possible that survey participants who report rounded point forecasts differ from respondents who round the probabilities. We show that this is the case below. Third, the employed

[18] The FED-SPF also conducted a special survey on the forecasting techniques of its participants in 2009Q4. 80% of the respondents reported that they use a mixture of a model-based approach and judgment. The summary does not specify whether the panelists were asked about the point forecasts, the histograms or their predictions in general.

[19] However, a discussion with one of the non-rounders from the ECB-SPF supports the notion of a close link between rounding and judgmental forecasting.

survey data differ in other important aspects. The sample used by Binder (2017) only covers a short period from January 2013 to September 2015, whereas we examine the SPF data for the period 1999Q1-2017Q4. Moreover, the bins in the SCE have a width of two percentage points and are thus much wider than those in the SPF. Furthermore, generalized beta distributions are fitted to the histograms of the SCE. Binder (2017) uses the interquartile range of the individual beta distributions in order to measure the dispersion. We follow Zarnowitz and Lambros (1987) and examine the individual variance as a measure of ex-ante uncertainty. Finally, the SCE differs from the SPF in terms of the sampling scheme by means of which surveyed individuals are selected. In particular, the SCE constitutes a rotating panel, whereas most of the SPF forecasters have a fairly long history of survey participation. The accumulated experience of some forecasters may also be related to their rounding choice.

In order to analyze whether the distinct approaches based on point and histogram forecasts isolate the same SPF participants, we first consider the correlations between the decimal- and integer-based categorizations for the reported probabilities based on a pooled sample of observations across all forecast horizons. Here, we only consider the case of the ECB-SPF. We have documented in the previous section that both approaches work well in isolating two distinct groups of forecasters who appear to rely on either judgment or models to compute their probabilities. If this is the case, the correlations between $D^{\text{any}}_{i,t,h}$ and $D^{\text{m5}}_{i,t,h}$ are expected to be positive and large.

In the second step, we follow Binder (2017) and categorize rounders based on whether the point prediction, $\mu^{\star}_{i,t,h}$, is a multiple of 0.5, i.e.,

$$\tilde{D}^{\text{m0.5}}_{i,t,h} = \begin{cases} 1 & \text{if } 0.5 \cdot \left\lfloor \frac{\mu^{\star}_{i,t,h}}{0.5} \right\rfloor = \mu^{\star}_{i,t,h} \text{ and} \\ 0 & \text{else.} \end{cases} \tag{5.24}$$

Note that Binder (2017) classifies consumers as rounders if the point forecast is a multiple of five, not 0.5. This is due to the fact that the range of point forecasts for inflation reported in the SCE is considerably larger than in the SPF. As in the case of the integer-based categorizations, we consider

$$D^{\text{m0.5}}_{i,t,h} = 1 - \tilde{D}^{\text{m0.5}}_{i,t,h} \tag{5.25}$$

in order to focus on non-rounders. If the categorizations based on point and histogram forecasts perform equally well, the correlations between $D^{\text{m0.5}}_{i,t,h}$ and either $D^{\text{any}}_{i,t,h}$ or $D^{\text{m5}}_{i,t,h}$ should also be positive and large. Table 5.3 summarizes the correlations based on a pooled sample of observations across all survey participants, time instances and forecast horizons.

Table 5.3: Correlations across categorizations in the ECB-SPF

	Inflation	GDP growth	Unemployment
$\widehat{\mathbf{Corr}}[D^{\text{any}}, D^{\text{m5}}]$	0.57	0.57	0.59
$\widehat{\mathbf{Corr}}[D^{\text{any}}, D^{\text{m0.5}}]$	−0.07	−0.07	−0.08
$\widehat{\mathbf{Corr}}[D^{\text{m5}}, D^{\text{m0.5}}]$	−0.06	−0.07	−0.08

Notes: For each outcome variable, this table displays the bivariate correlations between distinct categorizations for non-rounders in the ECB-SPF for a pooled sample of observations across all survey participants, time instances and forecast horizons. The sample period is 1999Q1–2016Q4.

The correlation statistics between $D^{\text{any}}_{i,t,h}$ and $D^{\text{m5}}_{i,t,h}$ have the expected sign and amount to 0.57, 0.57 and 0.59 for inflation, real GDP growth and unemployment, respectively. Thus, there is a large overlap in the groups of survey participants that are classified as non-rounders by both approaches. In contrast, the corresponding correlations between $D^{\text{any}}_{i,t,h}$ and $D^{\text{m0.5}}_{i,t,h}$ are considerably smaller and close to zero. In other words, these categorizations isolate distinct groups of forecasters. It may be the case that the weak association is due to methodological differences between the decimal-based approach and $D^{\text{m0.5}}_{i,t,h}$. If this were the only explanation, it may be expected that the categorizations from Eqns. (5.10) and (5.24) are more closely related, such that the association between $D^{\text{m5}}_{i,t,h}$ and $D^{\text{m0.5}}_{i,t,h}$ should be stronger. However, the corresponding correlation statistics are also close to zero for all variables. These results suggest that categorizations based on point or histogram forecasts isolate distinct groups of forecasters.

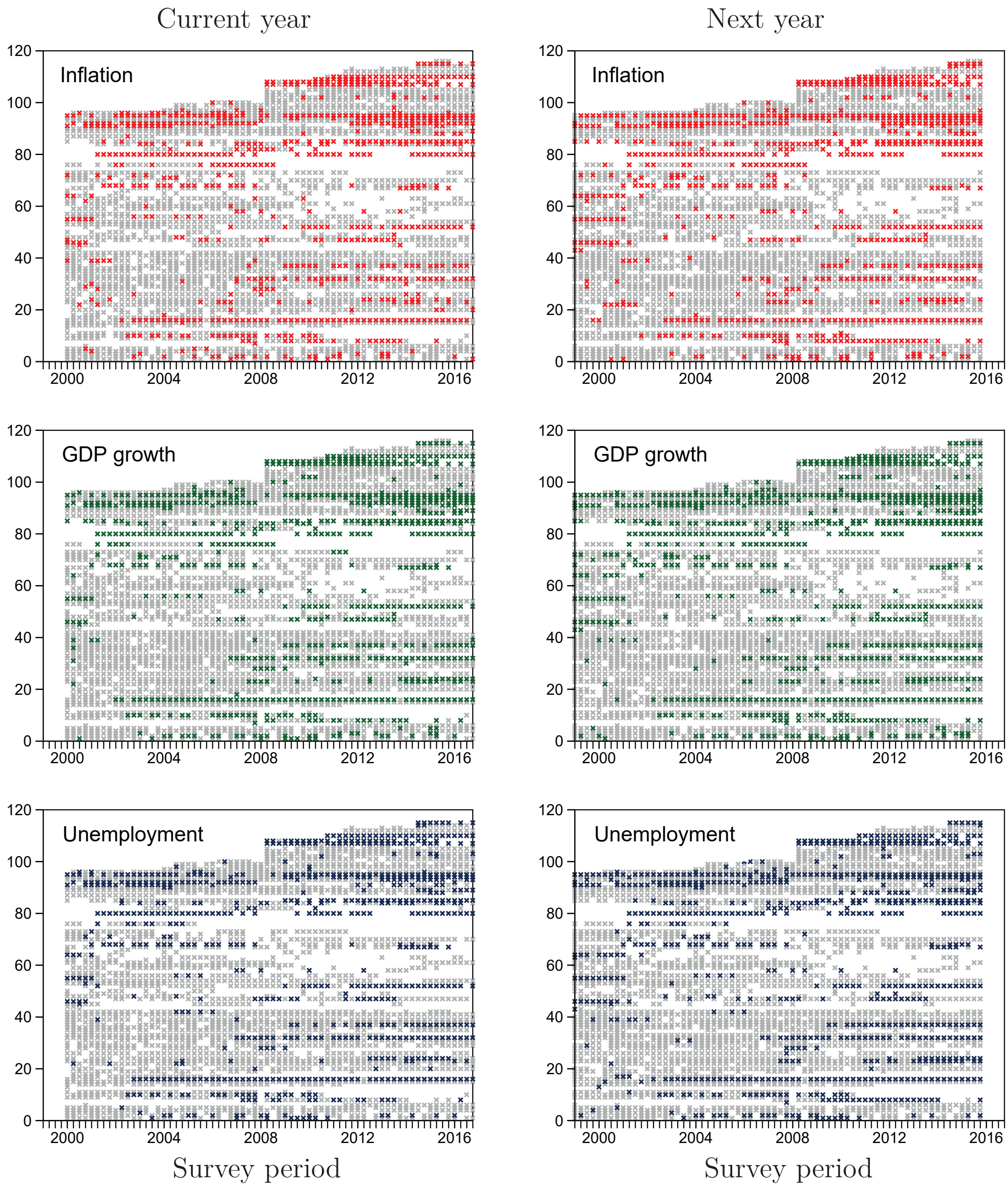
5.7 Conclusion

We analyze the misalignment between ex-ante and ex-post variances that is frequently observed in surveys of macroeconomic expectations. In the analysis of the Survey of Professional Forecasters for the Euro area and the U.S., we employ a variety of distinct categorizations to isolate two distinct types of forecasters based on their reporting behavior. We find that the variance misalignment is considerably smaller for survey participants who report non-rounded histogram forecasts. This is a consequence of the fact that this group reports significantly larger ex-ante variances. In contrast, the forecast errors of rounders and non-rounders do not seem to differ in a systematic way. Thus, rounding has little impact on the first-moment dynamics but has a substantial effect on the second moments. Our results have important implications for the evaluation of the cross-section of survey participants. In particular, more reliable measures of aggregate uncertainty can be derived by focusing on the non-rounders and discarding the remaining responses. Due to the relatively small share of non-rounded histograms, this would result in a substantial loss of information. An alternative solution could be the use of a weighting scheme where less weight is put on variances derived from predictions that are strongly rounded. Moreover, the share of non-rounders has increased over time. This suggests that the quality of the SPF predictions has improved in recent years and increases the feasibility of focusing on the non-rounders. Designers of surveys of macroeconomic expectations should improve their questionnaires in such a way that reporting less strongly rounded probabilities is further encouraged.

Our results also have implications for the reliability of using disagreement as a proxy for forecast uncertainty. Since we do not find evidence of substantial differences in the means of the histograms reported by rounders and non-rounders, measures of forecaster disagreement for both groups are likely to be relatively similar (see Binder, 2017). However, measures of aggregate uncertainty, e.g., the cross-sectional average variance, are strongly affected by the rounding choices of the panelists due the higher dispersion of the histograms reported by non-rounders. This suggests that one potential explanation for the increase in the difference between uncertainty and disagreement documented by, among others, Lahiri and Sheng (2010) and Glas and Hartmann (2016), is the growing share of non-rounders in recent survey periods.

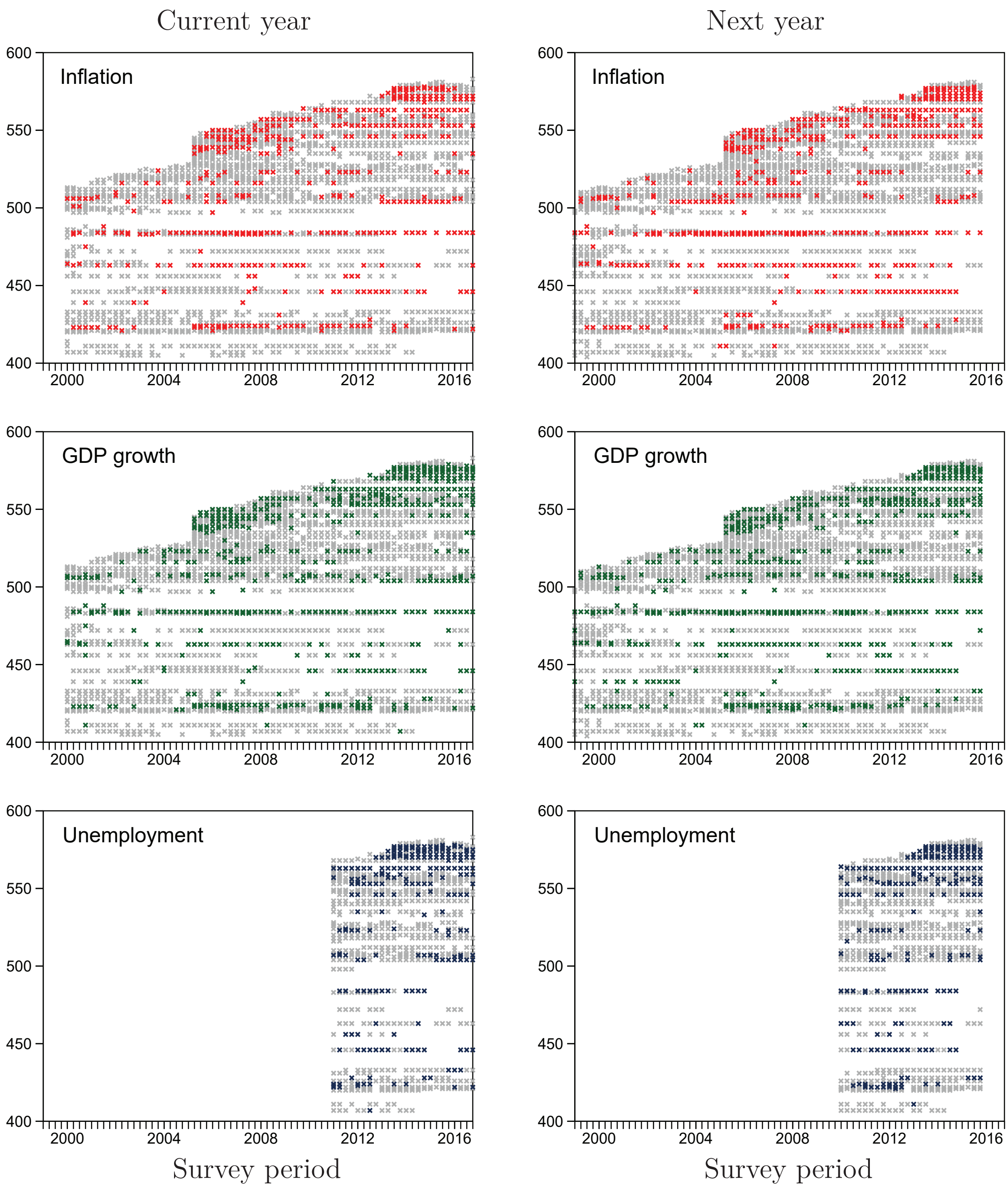
5.8 Appendix

Figure 5.16: Status of forecasters in the ECB-SPF



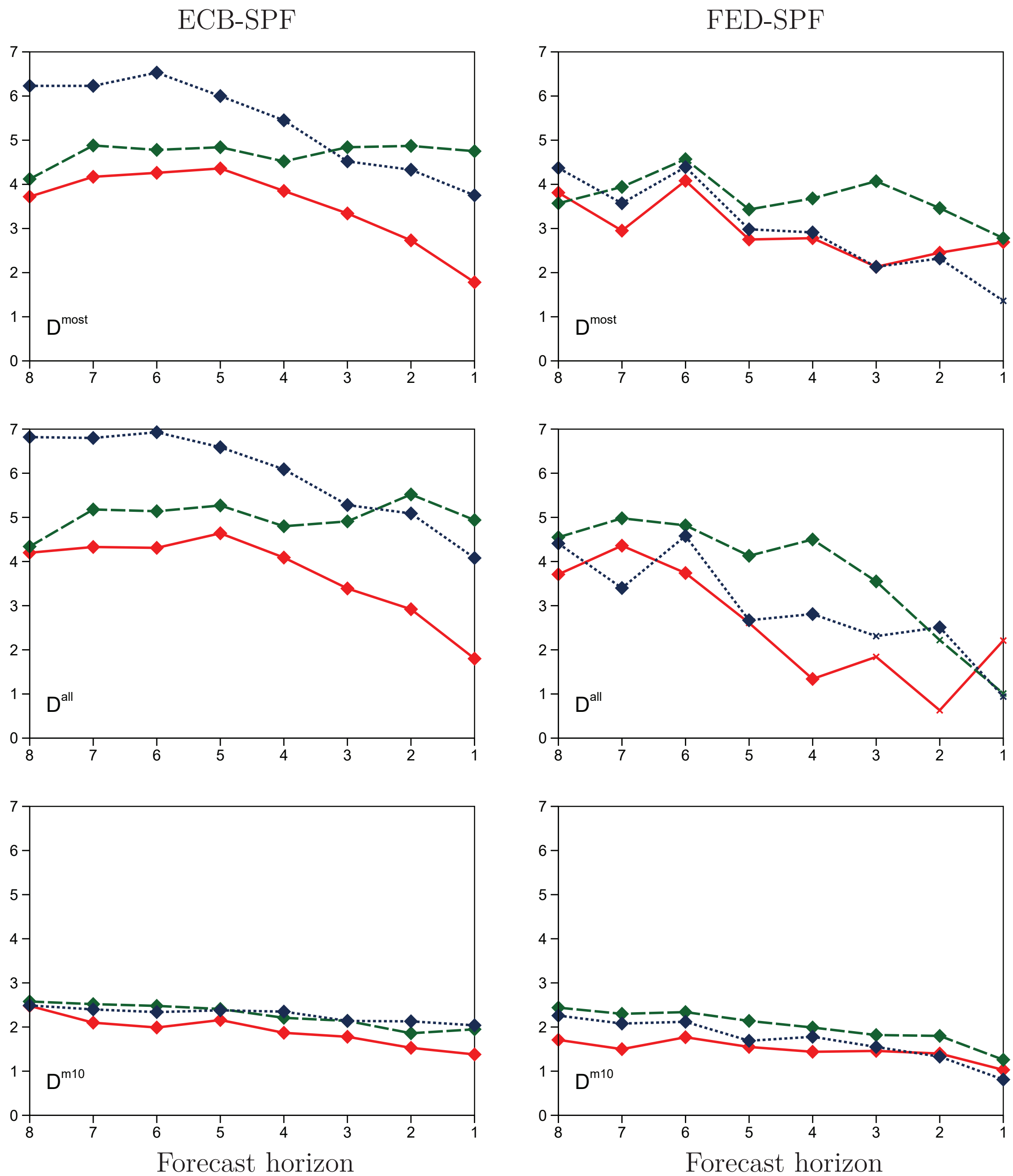
Notes: This figure depicts forecaster participation for the current year ($h \leq 4$) and next year ($h \geq 5$) forecasts in the ECB-SPF. The vertical axis indicates the identification number attached to individual forecasters. The horizontal axis depicts the period in which the predictions are reported. Each cross '$\times$' indicates that a histogram has been reported by a forecaster, i.e., cases where $D^{\mathcal{P}}_{i,t,h} = 1$. Non-rounders classified by means of $D^{\text{m5}}_{i,t,h}$ are highlighted as colored crosses. The sample period is 1999Q1–2016Q4.

Figure 5.17: Status of forecasters in the FED-SPF



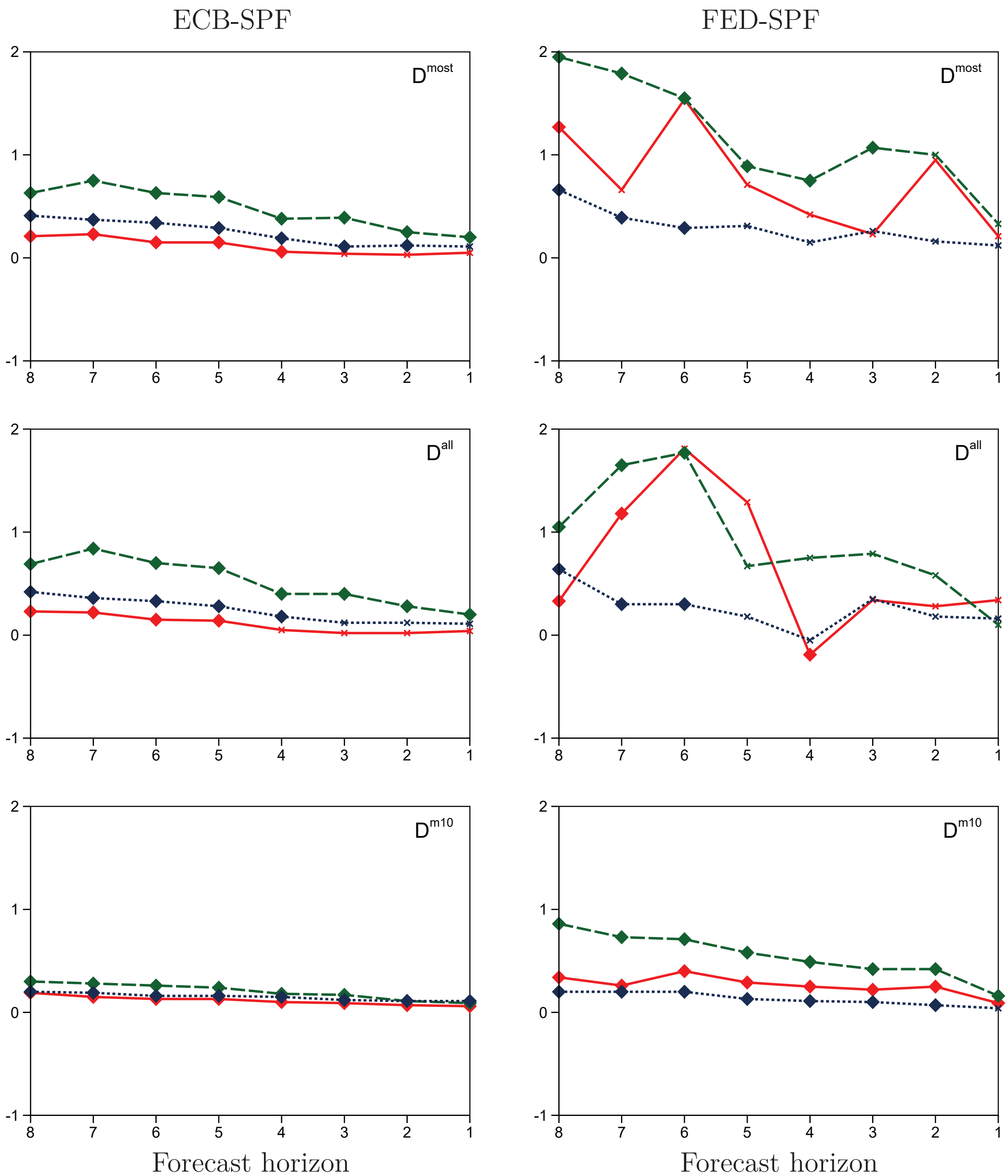
Notes: This figure depicts forecaster participation for the current year ($h \leq 4$) and next year ($h \geq 5$) forecasts in the FED-SPF. The vertical axis indicates the identification number attached to individual forecasters. The horizontal axis depicts the period in which the predictions are reported. Each cross '×' indicates that a histogram has been reported by a forecaster, i.e., cases where $D^{\mathcal{P}}_{i,t,h} = 1$. Non-rounders classified by means of $D^{\mathrm{m5}}_{i,t,h}$ are highlighted as colored crosses. The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts, which are available since 2010Q1 for our purposes. To improve the readability, we have excluded six forecasters with identification numbers below 100 from the graph.

Figure 5.18: Deviations in the number of bins used by non-rounders and rounders (other categorizations)



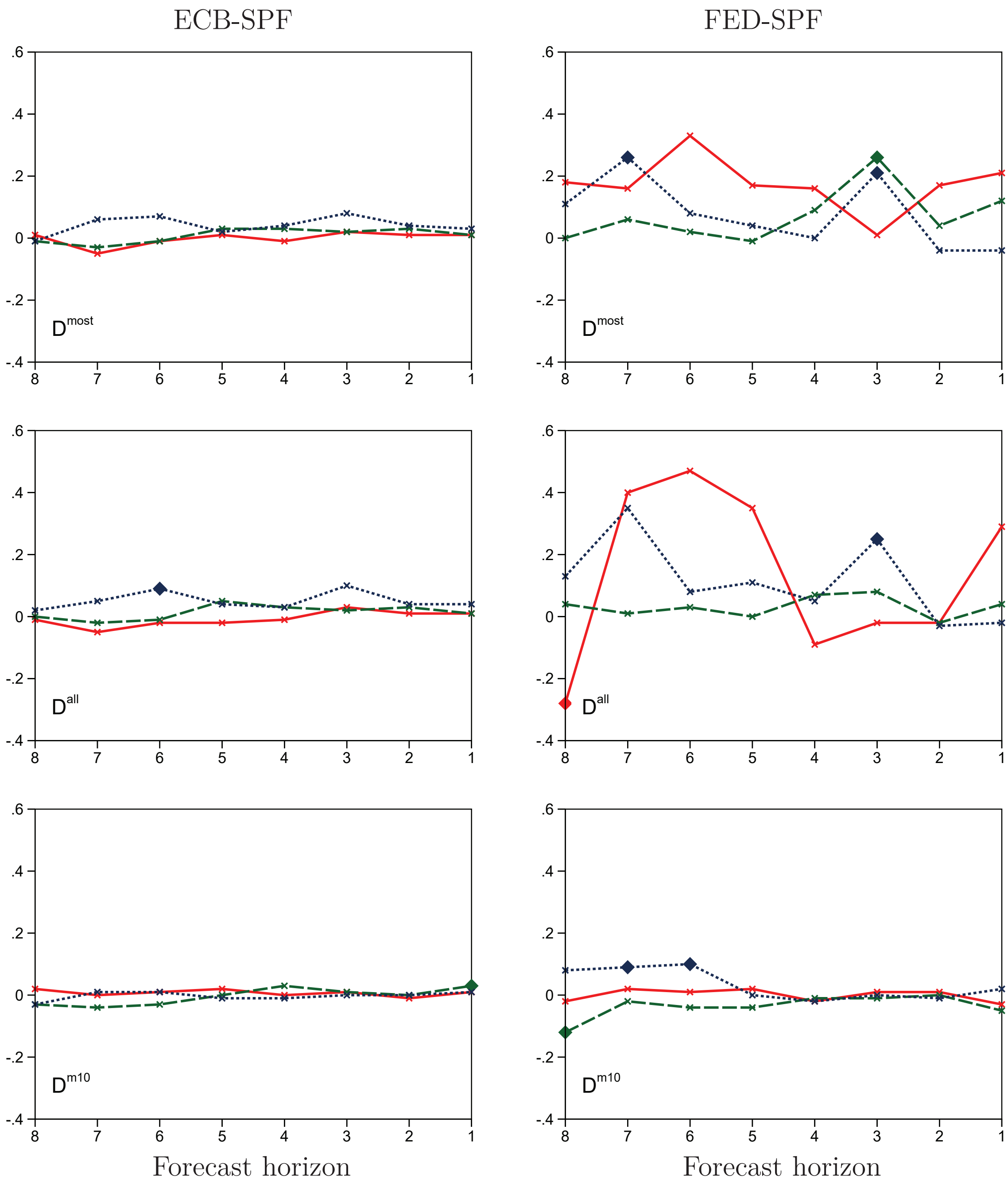
Notes: For each forecast horizon, the graphs depict the difference in the number of bins used by non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $K_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the number of bins used is distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4 that are not considered in the main text. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.19: Deviations in the variances reported by non-rounders and rounders (other categorizations)



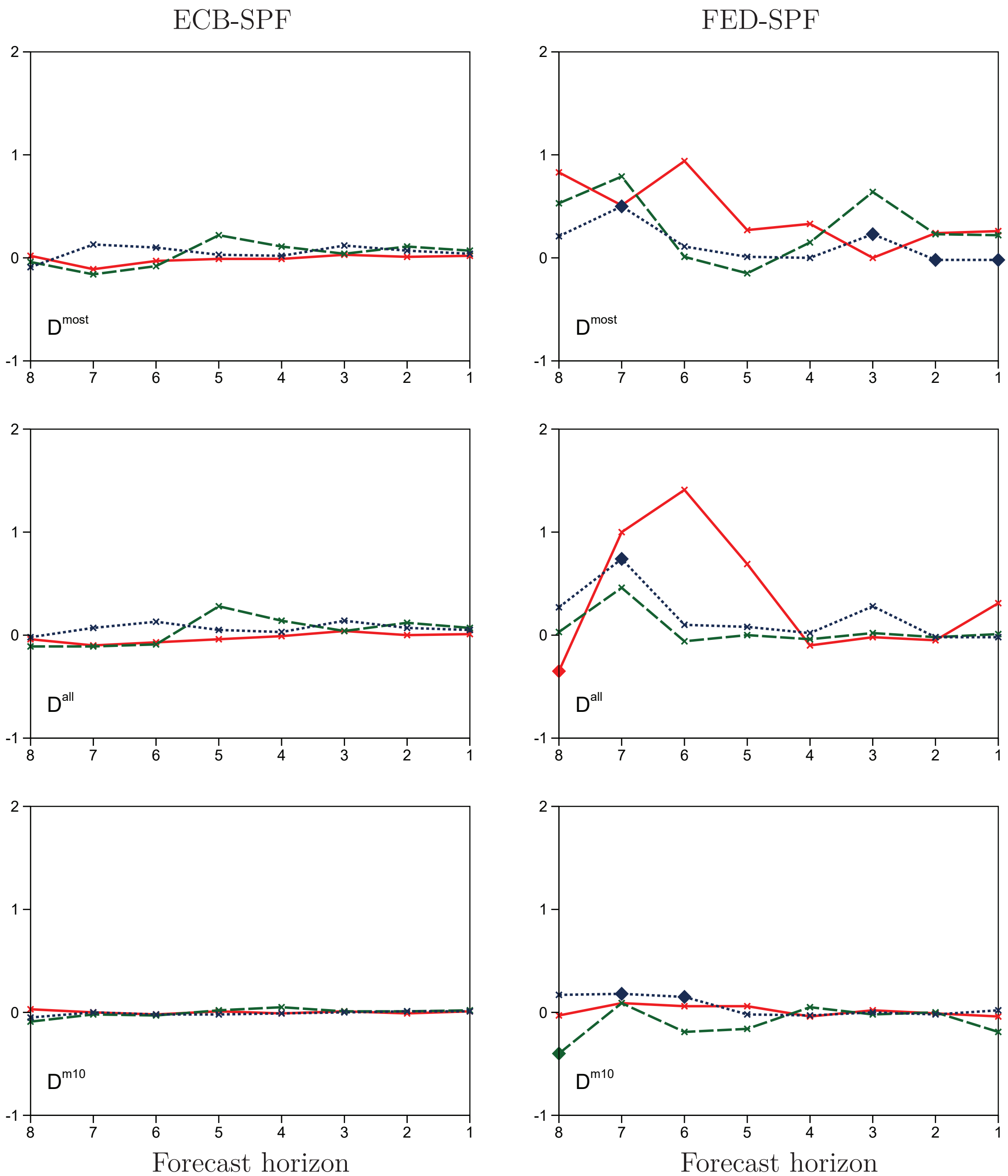
Notes: For each forecast horizon, the graphs depict the difference in the ex-ante variances reported by non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $\sigma^2_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the reported variance forecasts are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4 that are not considered in the main text. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.20: Deviations in the absolute forecast errors of non-rounders and rounders (other categorizations)



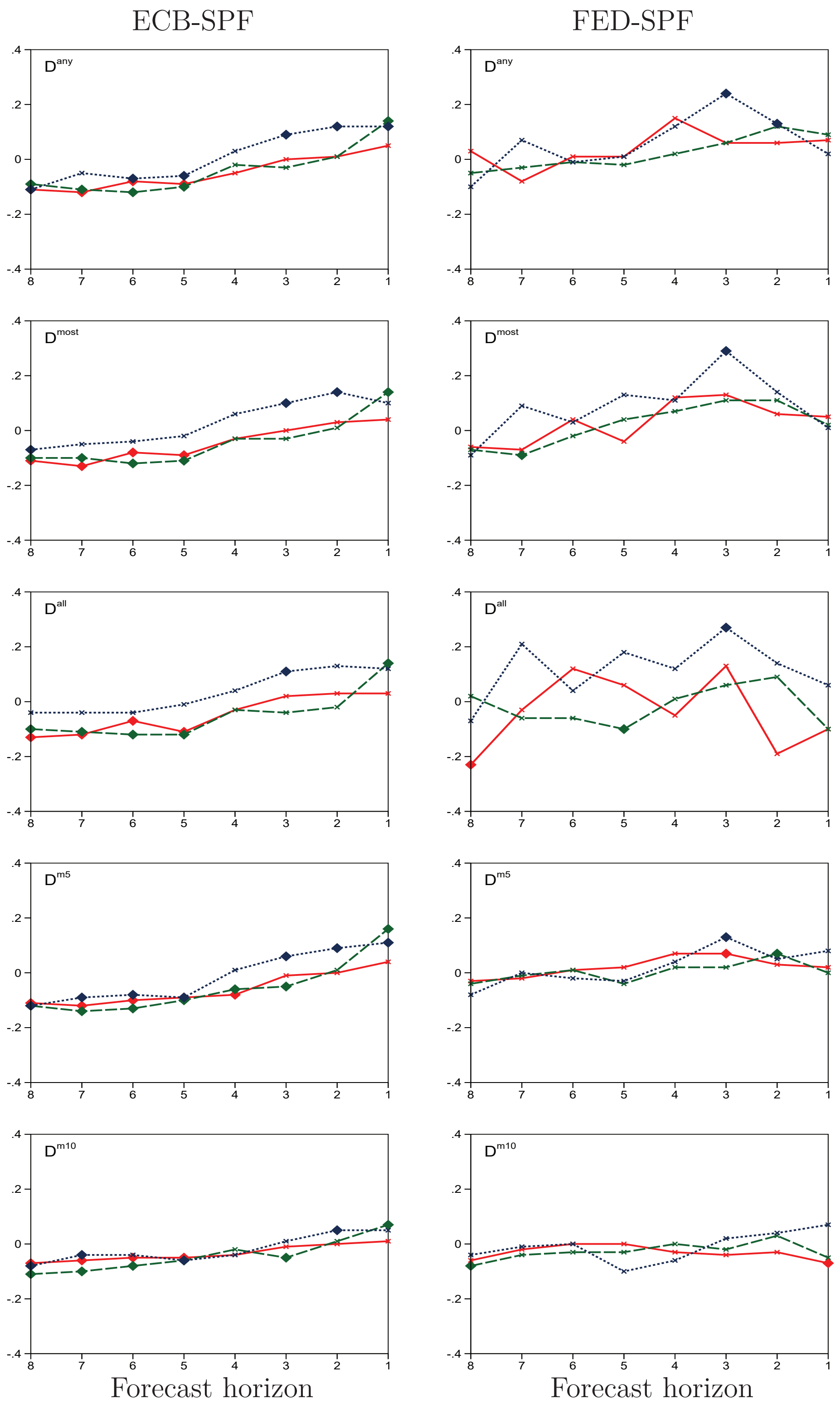
Notes: For each forecast horizon, the graphs depict the difference in the ex-post absolute forecast errors of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $|e_{i,t,h}|$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the absolute prediction errors are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4 that are not considered in the main text. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.21: Deviations in the squared forecast errors of non-rounders and rounders (other categorizations)



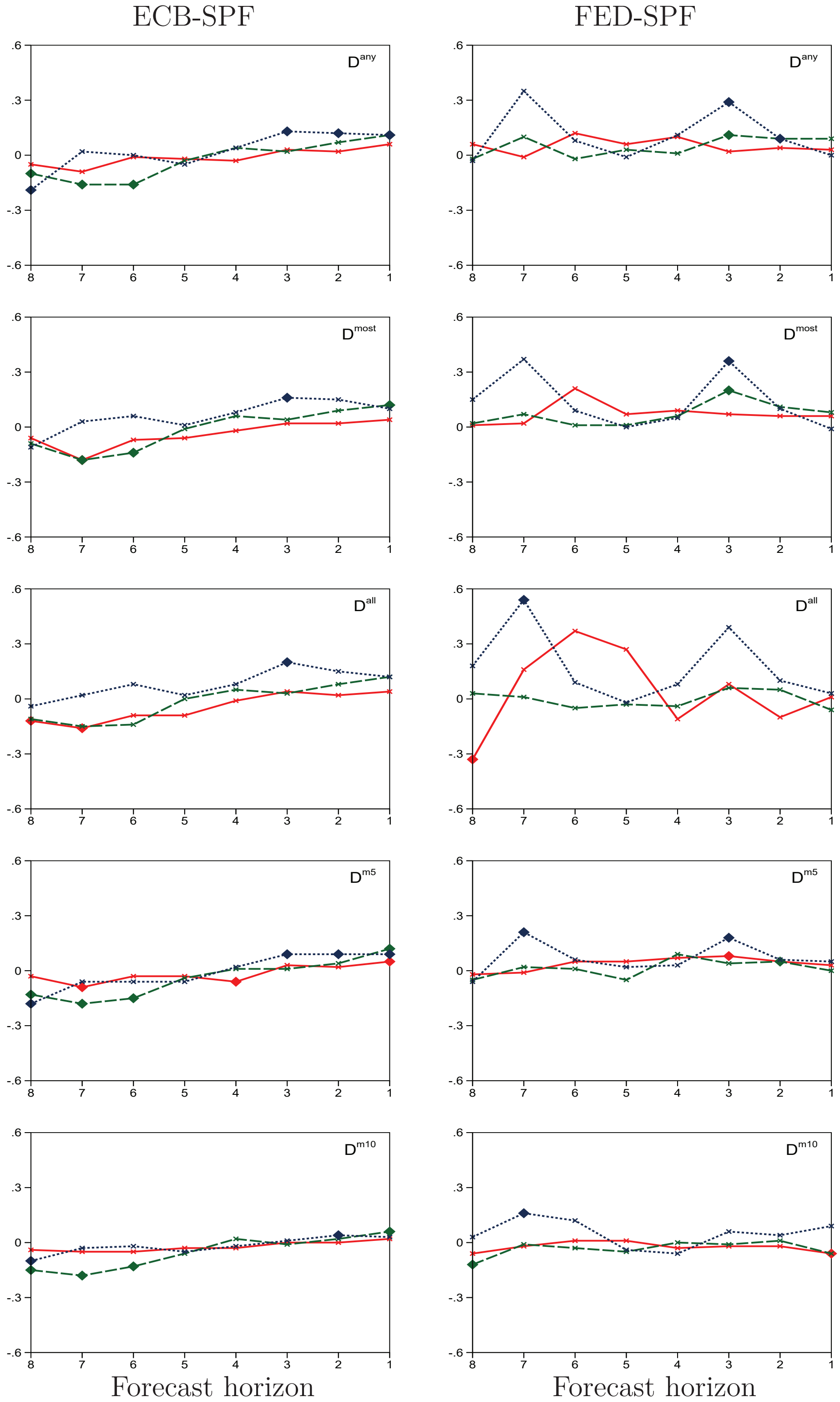
Notes: For each forecast horizon, the graphs depict the difference in the ex-post squared forecast errors of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $e^2_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the squared prediction errors are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4 that are not considered in the main text. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.22: Deviations in the QPS scores of non-rounders and rounders



Notes: For each forecast horizon, the graphs depict the difference in the quadratic probability scores of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $\text{QPS}_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the QPS scores are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m5}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Figure 5.23: Deviations in the RPS scores of non-rounders and rounders



Notes: For each forecast horizon, the graphs depict the difference in the ranked probability scores of non-rounders and rounders. In particular, each marker denotes an estimate of the slope coefficient, β_h, based on the h-step-ahead predictions for inflation (solid), output growth (dashed) and unemployment (dotted) when $\text{RPS}_{i,t,h}$ is considered as the dependent variable in Eqn. (5.23). A diamond '$\diamond$' indicates that the RPS scores are distinct among non-rounders and rounders. The significance level is 5%. A cross '$\times$' indicates an insignificant estimate. The explanatory variable $D^{\mathcal{R}}_{i,t,h} \in \{D^{\text{any}}_{i,t,h}, D^{\text{most}}_{i,t,h}, D^{\text{all}}_{i,t,h}, D^{\text{m5}}_{i,t,h}, D^{\text{m10}}_{i,t,h}\}$ denotes one of the decimal- or integer-based categorizations from Section 5.4. Each regression includes time-fixed effects. Coefficients are estimated via OLS. We apply the variance-covariance estimator of Newey and West (1987). The sample period is 1999Q1–2016Q4, except for the unemployment rate forecasts from the FED-SPF, which are available since 2010Q1 for our purposes.

Bibliography

[1] Abel, J., Rich, R., Song, J., Tracy, J., 2016. The measurement and behavior of uncertainty: Evidence from the ECB Survey of Professional Forecasters. *Journal of Applied Econometrics* 31, 533–550.

[2] Andersen, T.G., Bollerslev, T., Christoffersen, P.F., Diebold, F.X., 2013. Financial risk measurement for financial risk management. In: *Handbook of the Economics of Finance* Vol. II (eds. Constanides, G., Harris, M., Stulz, R.), Chapter 17, pp. 1127–1220. Amsterdam: Elsevier Science B.V.

[3] Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics* 73, 251–277.

[4] Andrade, P., Crump, R.K., Eusepi, S., Moench, E., 2016. Fundamental disagreement. *Journal of Monetary Economics* 83, 106–128.

[5] Andrade, P., Ghysels, E., Idier, J., 2015. Tails of inflation forecasts and tales of monetary policy. Banque de France WP-407.

[6] Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54, 1163–1212.

[7] Asgharian, H., Hess, W., Liu, L., 2013. A spatial analysis of international stock market linkages. *Journal of Banking & Finance* 37, 4738–4754.

[8] Bachmann, R., Elstner, S., Sims, E. R., 2013. Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics* 5, 217–249.

[9] Baillie, R.T., Chung, C.-F., Tieslau, M.A., 1996. Analysing inflation by the fractionally integrated ARIMA-GARCH model. *Journal of Applied Econometrics* 11, 23–40.

[10] Baillie, R.T., Han, Y.W., Kwon, T.-G., 2002. Further long memory properties of inflationary shocks. *Southern Economic Journal* 68, 496–510.

[11] Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. *The Quarterly Journal of Economics* 131, 1593–1636.

[12] Ball, L., 1992. Why does high inflation raise inflation uncertainty? *Journal of Monetary Economics* 29, 371–388.

[13] Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59, 1481–1509.

[14] Barron, O.E., Kim, O., Lim, S.C., Stevens, D.E., 1998. Using analysts' forecasts to measure properties of of analysts' information environment. *Accounting Review* 73, 421–433.

[15] Barsky, R.B., Sims, E.R., 2012. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review* 102, 1343–1377.

[16] Binder, C.C., 2017. Measuring uncertainty based on rounding: New method and application to inflation expectations. *Journal of Monetary Economics* 90, 1–12.

[17] Bloom, N., 2009. The impact of uncertainty shocks. *Econometrica* 77, 623–685.

[18] Bloom, N., Fernández-Villaverde, J., Schneider, M., 2014. The macroeconomics of uncertainty and volatility. *Journal of Economic Literature*, forthcoming.

[19] Blundell, R., Bond, S., 2013. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115–143.

[20] Boero, G., Smith, J., Wallis, K.F., 2008. Uncertainty and disagreement in economic prediction: The Bank of England Survey of External Forecasters. *The Economic Journal* 118, 1107–1127.

[21] Boero, G., Smith, J., Wallis, K.F., 2011. Scoring rules and survey density forecasts. *International Journal of Forecasting* 27, 379–393.

[22] Boero, G., Smith, J., Wallis, K.F., 2015. The measurement and characteristics of professional forecasters' uncertainty. *Journal of Applied Econometrics* 30, 1029–1046.

[23] Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. *Review of Financial Studies* 22, 4463–4492.

[24] Boyd, J.H., Hu, J., Jagannathan, R., 2005. The stock market's reaction to unemployment news: Why bad news is usually good for stocks. *The Journal of Finance* 60, 649–672.

[25] Branch, W.A., 2004. The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *The Economic Journal* 114, 592–621.

[26] Bruine De Bruin, W., Carman, K.G., 2012. What does the excessive use of 50% mean? *Medical Decision Making* 32, 232–236.

[27] Bruine de Bruin, W., Potter, S., Rich, R., Topa, G., van der Klaauw, W., August/September 2010. Improving survey measures of household inflation expectations. *Current Issues in Economics and Finance* 16, 1–7.

[28] Bruine De Bruin, W., van der Klaauw, W., Topa, G., 2011. Expectations of inflation: The biasing effect of thoughts about specific prices. *Journal of Economic Psychology* 32, 834–845.

[29] Bruine De Bruin, W., van der Klaauw, W., Topa, G., Downs, J.S., Fischhoff, B., Armantier, O. 2012. The effect of question wording on consumers' reported inflation expectations. *Journal of Economic Psychology* 33, 749–757.

[30] Bry, G., Boschan, C., 1971. Cyclical analysis of time series: Selected procedures and computer programs. New York, NBER.

[31] Campbell, J.Y., 1991. A variance decomposition for stock returns. *Economic Journal* 101, 157–179.

[32] Campbell, J.Y., Giglio, S., Polk, C., Turley, R., 2018. An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics* 128, 207–233.

[33] Campbell, J.Y., Hentschel, L., 1992. No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Econometrics* 31, 281–318.

[34] Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance* 56, 1–43.

[35] Campbell, J.Y., Shiller, R.J., 1988. Stock prices, earnings, and expected dividends. *The Journal of Finance* 43, 661–676.

[36] Campbell, J.Y., Shiller, R.J., 2001. Valuation ratios and the long-run stock market outlook: An update. Cowles Foundation Discussion Paper No. 1295.

[37] Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies* 21, 1509–1531.

[38] Campbell, S.D., Diebold, F.X., 2009. Stocks returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics* 27, 266–278.

[39] Capistran, C., Timmermann, A., 2009. Forecast combination with entry and exit of experts. *Journal of Business & Economic Statistics* 27, 428–440.

[40] Chauvet, M., Hamilton, J.D., 2005. Dating business cycles turning points. NBER Working Paper 11422.

[41] Chong, T.T.-L., Lin, S., 2017. Predictive models for disaggregate stock market volatility. *Financial Markets and Portfolio Management* 31, 261–288.

[42] Christiansen, C., Schmeling, M., Schrimpf, A., 2012. A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics* 27, 956–977.

[43] Clarida, R., Galí, J., Gertler, M., 1998. Monetary policy rules in practice: Some international evidence. *European Economic Review* 42, 1033–1067.

[44] Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291–311.

[45] Clements, M.P., 2010. Explanations of inconsistencies in survey respondent's forecasts. *European Economic Review* 54, 536–549.

[46] Clements, M.P., 2011. An empirical investigation of the effects of rounding on the SPF probabilities of decline and output growth histograms. *Journal of Money, Credit and Banking* 43, 207–220.

[47] Clements, M.P., 2012. Probability distributions or point predictions? Survey forecasts of US output growth and inflation. Warwick Economic Research Papers 976, 1–36.

[48] Clements, M.P., 2014. Forecast uncertainty – ex ante and ex-post: U.S. inflation and output growth. *Journal of Business & Economic Statistics* 32, 206–216.

[49] Clements, M.P., 2014. Probability distributions or point predictions? Survey forecasts of US output growth and inflation. *International Journal of Forecasting* 30, 99–117.

[50] Clements, M.P., 2016. Are macroeconomic density forecasts informative? ICMA Centre Discussion Papers in Finance No. DP2016-02.

[51] Cochrane, J., 1992. Explaining the variance of price-dividend ratios. *Review of Financial Studies* 5, 243–280.

[52] Conrad, C., Glas, A., 2018. 'Déjà vol' revisited: Survey forecasts of macroeconomic variables predict volatility in the cross-section of industry portfolios. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3186567` (last accessed: 27 July 2018).

[53] Conrad, C., Hartmann, M., 2014. Cross-sectional evidence on the relation between monetary policy, macroeconomic conditions and low-frequency inflation uncertainty. University of Heidelberg Discussion Paper No. 574.

[54] Conrad, C., Karanasos, M., 2005. Dual long memory in inflation dynamics across countries of the Euro area and the link between inflation uncertainty and macroeconomic performance. *Studies in Nonlinear Dynamics & Econometrics* 9, Article 5.

[55] Conrad, C., Karanasos, M., Zeng, N., 2010. The link between macroeconomic performance and variability in the UK. *Economics Letters* 106, 154–157.

[56] Conrad, C., Kleen, O., 2018. Two are better than one: Volatility forecasting using multiplicative component GARCH models. Available at SSRN: `https://ssrn.com/abstract=2752354` (last accessed: 27 July 2018).

[57] Conrad, C., Lamla, M.J., 2010. The high-frequency response of the EUR-USD exchange rate to ECB communication. *Journal of Money, Credit and Banking* 42, 1391–1417.

[58] Conrad, C., Loch, K., 2015. Anticipating long-term stock market volatility. *Journal of Applied Econometrics* 30, 1090–1114.

[59] Conrad, C., Schienle, M., 2018. Testing for an omitted multiplicative long-term component in GARCH models. *Journal of Business & Economic Statistics*, forthcoming.

[60] Cosimano, T.F., Jansen, D.W., 1988. Estimates of the variance of U.S. inflation based upon the ARCH model: Comment. *Journal of Money, Credit and Banking* 20, 409–421.

[61] D'Amico, S., Orphanides, A., 2008. Uncertainty and disagreement in economic forecasting. Finance and Economics Discussion Series. Board of Governors of the Federal Reserve System 2008-56.

[62] David, A., Veronesi, P., 2013. What ties return volatilities to price valuations and fundamentals? *Journal of Political Economy* 121, 682–746.

[63] Dawid, A.P., 1984. Statistical theory: The prequential approach (with discussion). *Journal of the Royal Statistical Society: Series A* 147, 278–292.

[64] Dovern, J., 2015. A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data. *European Economic Review* 80, 16–35.

[65] Dovern, J., Fritsche, U., Slacalek, J., 2012. Disagreement among forecasters in G7 countries. *Review of Economics and Statistics* 94, 1081–1096.

[66] Drakos, A.A., Kouretas, G.P., 2015. The conduct of monetary policy in the Eurozone before and after the financial crisis. *Economic Modelling* 48, 83–92.

[67] Dräger, L., Lamla, M.J., Pfajfar, D., 2016. Are survey expectations theory-consistent? The role of central bank communication and news. *European Economic Review* 85, 84–111.

[68] Engel, C., West, K.D., 2006. Taylor rules and the Deutschmark: Dollar real exchange rate. *Journal of Money, Credit and Banking* 38, 1175–1194.

[69] Engelberg, J., Manski, C.F., Williams, J., 2009. Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business & Economic Statistics* 27, 30–41.

[70] Engle, R.F., 1983. Estimates of the variance of U.S. inflation based upon the ARCH model. *Journal of Money, Credit and Banking* 15, 286–301.

[71] Engle, R.F., Ghyssels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics* 95, 776–797.

[72] Engle, R.F., Rangel, J.G., 2008. The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies* 21, 1187–1222.

[73] Engle, R.F., Rangel, J.G., 2012. The Factor-Spline-GARCH model for high and low frequency correlations. *Journal of Business & Economic Statistics* 30, 109–124.

[74] Estrella, A., Mishkin, F.S., 1997. The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review* 41, 1375–1401.

[75] Estrella, A., Mishkin, F.S., 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* 80, 45–61.

[76] European Central Bank, 2009. Results of a special questionnaire for participants in the ECB Survey of Professional Forecasters. Available at: `https://www.ecb.europa.eu/stats/prices/indic/forecast/shared/files/quest_summary.pdf` (last accessed: 27 July 2018).

[77] European Central Bank, 2014a. ECB Monthly Bulletin November 2014.

[78] European Central Bank, 2014b. Results of the second special questionnaire for participants in the ECB Survey of Professional Forecasters. Available at: `https://www.ecb.europa.eu/stats/prices/indic/forecast/shared/files/resultssecondspecialquestionnaireecbsurveyspf201401en.pdf` (last accessed: 27 July 2018).

[79] Fama, E.F., 1981. Stock returns, real activity, inflation, and money. *American Economic Review* 71, 545–565.

[80] Fama, E.F., Schwert, G.W., 1977. Asset returns and inflation. *Journal of Financial Economics* 5, 115–146.

[81] Faust, J., Wright, J.H., 2009. Comparing Greenbook and reduced form forecasts using a large realtime dataset. *Journal of Business & Economic Statistics* 27, 468–479.

[82] Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J.F., 2011. Risk matters: The real effects of volatility shocks. *American Economic Review* 101, 2530–2561.

[83] Friedman, M., 1977. Inflation and unemployment. *Journal of Political Economy* 85, 451–472.

[84] Galí, J., Gambetti, L., 2009. On the sources of the Great Moderation. *American Economic Journal: Macroeconomics* 1, 26–57.

[85] Giacomini, R., White, H., 2006. Tests of conditional predictive ability. *Econometrica* 74, 1545–1578.

[86] Giordani, P., Söderlind, P., 2003. Inflation forecast uncertainty. *European Economic Review* 47, 1037–1059.

[87] Giordani, P., Söderlind, P., 2006. Is there evidence of pessimism and doubt in subjective distributions? Implications for the equity premium puzzle. *Journal of Economic Dynamics and Control* 30, 1027–1043.

[88] Glas, A., 2018. Five dimensions of the uncertainty-disagreement linkage. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3295125` (last accessed: 28 December 2018).

[89] Glas, A., Hartmann, M., 2016. Inflation uncertainty, disagreement and monetary policy: Evidence from the ECB Survey of Professional Forecasters. *Journal of Empirical Finance* 39, 215–228.

[90] Glas, A., Hartmann, M., 2018. Overconfidence versus rounding in survey-based density forecasts. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3202810` (last accessed: 27 July 2018).

[91] Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21, 1455–1508.

[92] Grier, K.B., Henry, O.T., Olekalns, N., Shields, K., 2004. The asymmetric effects of uncertainty on inflation and output growth. *Journal of Applied Econometrics* 19, 551–565.

[93] Grier, K.B., Perry, M.J., 2000. The effects of real and nominal uncertainty on inflation and output growth. *Journal of Applied Econometrics* 15, 45–58.

[94] Grimme, C., Henzel, S.R., Wieland, E., 2014. Inflation uncertainty revisited: A proposal for robust measurement. *Empirical Economics* 47, 1497–1523.

[95] Harding, D., Pagan, A., 2002. Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics* 49, 365–381.

[96] Hartmann, M., Herwartz, H., 2012. Causal relations between inflation and inflation uncertainty – Cross-sectional evidence in favor of the Friedman-Ball hypothesis. *Economics Letters* 115, 144–147.

[97] Hartmann, M., Roestel, J., 2013. Inflation, output and uncertainty in the era of inflation targeting – A multi-economy view on causal linkages. *Journal of International Money and Finance* 37, 98–112.

[98] Heitjan, D.F., Rubin, D.B., 1991. Ignorability and coarse data. *The Annals of Statistics* 19, 2244–2253.

[99] Henzel, S.R., Rengel, M., 2017. Dimensions of macroeconomic uncertainty: A common factor analysis. *Economic Inquiry* 55, 843–877.

[100] Holland, A.S., 1995. Inflation and uncertainty-tests for temporal ordering. *Journal of Money, Credit and Banking* 27, 827–837.

[101] Jo, S., Sekkel, R., 2018. Macroeconomic uncertainty through the lens of professional forecasters. *Journal of Business & Economic Statistics*, forthcoming.

[102] Jurado, K., Ludvigson, S.C., Ng, S., 2015. Measuring uncertainty. *American Economic Review* 105, 1177–1216.

[103] Kenny, G., Kostka, T., Masera, F., 2014. How informative are the subjective density forecasts of macroeconomists? *Journal of Forecasting* 33, 163–185.

[104] Kenny, G., Kostka, T., Masera, F., 2015. Density characteristics and density forecast performance: A panel analysis. *Empirical Economics* 48, 1203–1231.

[105] Klein, B., 1977. The demand for quality-adjusted cash balances: Price uncertainty in the U.S. demand for money function. *Journal of Political Economy* 85, 691–715.

[106] Krüger, F., 2017. Survey-based forecast distributions for Euro Area growth and inflation: Ensembles versus histograms. *Empirical Economics* 53, 235–246.

[107] Lahiri, K., Liu, F., 2006. Modelling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics* 21, 1199–1219.

[108] Lahiri, K., Peng, H., Sheng, X., 2015. Measuring uncertainty of a combined forecast and some tests for forecaster heterogeneity. CESIFO Working Paper No. 5468, 1–52.

[109] Lahiri, K., Peng, H., Zhao, Y., 2017. Online learning and forecast combination in unbalanced panels. *Econometric Reviews* 36, 257–288.

[110] Lahiri, K., Sheng, X., 2008. Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics* 144, 325–340.

[111] Lahiri, K., Sheng, X., 2010. Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics* 25, 514–538.

[112] Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance* 56, 815–849.

[113] López-Pérez, V., 2016. Does uncertainty affect non-response to the European Central Bank's Survey of Professional Forecasters? *Economics E-Journal* 10, 1–46.

[114] Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123, 137–162.

[115] Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement about inflation expectations. NBER Working Paper 9796.

[116] Manski, C.F., 2004. Measuring expectations. *Econometrica* 72, 1329–1376.

[117] Manski, C.F., Molinari, F., 2010. Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics* 28, 219–231.

[118] Meinen, P., Röhe, O., 2017. On measuring uncertainty and its impact on investment: Cross-country evidence from the Euro area. *European Economic Review* 92, 161–179.

[119] Mele, A., 2007. Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics* 86, 446–478.

[120] Mittnik, S., Robinzonov, N., Spindler, M., 2015. Stock market volatility: Identifying major drivers and the nature of their impact. *Journal of Banking & Finance* 58, 1–14.

[121] Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.

[122] Officer, R.R., 1973. The variability of the market factor of New York Stock Exchange. *Journal of Business* 46, 434–453.

[123] Pástor, L., Veronesi, P., 2013. Policy uncertainty and risk premia. *Journal of Financial Economics* 110, 520-545.

[124] Paye, B.S., 2012. 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics* 106, 527–546.

[125] Poncela P., Senra, E., 2017. Measuring uncertainty and assessing its predictive power in the Euro area. *Empirical Economics* 53, 165–182.

[126] Rich, R., Tracy, J., 2010. The relationships among expected inflation, disagreement, and uncertainty: Evidence from matched point and density forecasts. *The Review of Economics and Statistics* 92, 200–207.

[127] Rossi, B., Sekhposyan, T., 2017. Macroeconomic uncertainty indices for the Euro area and its individual member countries. *Empirical Economics* 53, 41–62.

[128] Ruud, P.A., Schunk, D., Winter, J.K., 2014. Uncertainty causes rounding: An experimental study. *Experimental Economics* 17, 391–413.

[129] Sauter, O., 2012. Assessing uncertainty in Europe and the US – Is there a common factor? FZID Discussion Papers No. 47-2012.

[130] Schwert, G.W., 1989. Why does stock market volatility change over time? *The Journal of Finance* 44, 1115–1153.

[131] Söderlind, P., 2011. Inflation risk premia and survey evidence on macroeconomic uncertainty. *International Journal of Central Banking* 7, 113–133.

[132] Stock, J.H., Watson, M.W., 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 3–33.

[133] Tay, A.S., Wallis, K.F., 2002. Density forecasting: A survey. *Companion to Economic Forecasting*, 45–68.

[134] Taylor, J.B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.

[135] Taylor, J.B., 2012. Monetary policy rules work and discretion doesn't: A tale of two eras. *Journal of Money, Credit and Banking* 44, 1017–1032.

[136] Windmeijer, F. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126, 25–51.

[137] Wright, J.H., 2011. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. *American Economic Review* 101, 1514–1534.

[138] Zarnowitz, V., Lambros, L.A., 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95, 591–621.